Part 5- Session Presentations for the EPA 23rd Annual National Conference on Managing Environmental Quality Systems

April 13-16, 2004 Tampa, Florida

Ambient Air I

Does Pulsed Sampling Provide Good Estimates of Atmospheric Concentrations? (Roger Carter and Debbie Lacroix, NOAA)

National Air Toxics Trends Stations Quality System (Dennis Mikel and Danny France, U.S. EPA)

Ambient Air II

Contributions of Tribal Environmental Agencies to Quality Assurance in Environmental Monitoring (Melinda Ronca-Battista, Northern Arizona University)

Multi-Site Evaluations of Candidate Methodologies for Determiningg Coarse particulate Matter (PMc) Concentrations (Robert Vanderpool, U.S. EPA and Andrew Johnson, Maine DEP)

Ambient Air III

Data Quality Objective Development for the Coarse Particualte Matter (PMc) Standard (Mike Papp, U.S. EPA)

New Quality Indicator Statistics for the Gaseous Criteria Pollutants (Basil Coutant, Battlelle)

Status and Changes in EPA Infrastructure for Bias Traceability to NIST (Mark Shanis, U.S. EPA)

Ambient Air IV

2003 EPA Audit of Protocol Cylinder Gases (Joe Elkins and B. Wright, U.S. EPA)

NPEP Through the Probe Demonstration (Greg Noah, U.S. EPA)

Does Pulsed Sampling Provide Good Estimates of Atmospheric Concentrations?

Roger G. Carter and Debbie J. Lacroix
National Oceanic and Atmospheric Administration
Air Resoures Laboratory
Field Research Division (NOAA-ARLFRD)

Tracer Experiments

- To understand how the atmosphere transports, disperses, and diffuses materials.
  - Release a small amount of tracer.
  - Sample and measure the concentrations.
  - Combine results with meteorological information.

How is Sampling Done?

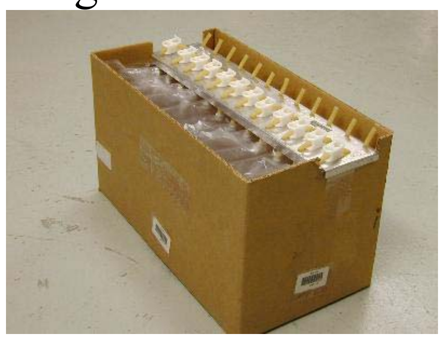
- Programmable Integrating Gas Sampler
  - Collect samples over a period.
  - Sequentially fill 12 bags for the programmed duration.
  - Pumps are “pulsed” for short bursts.

Sampler

- Waxed cardboard box with pumps and controlling electronics.

Cartridge

- Smaller cardboard box containing 12 Tedlar bags.



Sampler + Cartridge Inside Without the Lid

- The cartridge is placed inside the sampler.

Sample + Cartridge With Lid

- The lid is placed on the sampler.

How Are the PIGS Placed?

- PIGS are placed at pre-selected sites.
- Hung 10 feet above ground on poles.

Keeping Track of All Those Samples

- Locations, samplers and cartridges tagged with barcodes.
- Longitude and latitude recorded on a computer.
- Timewands used to track samples.

Timewand

Automated Tracer Gas Analysis System (ATGAS)

Sample Analysis

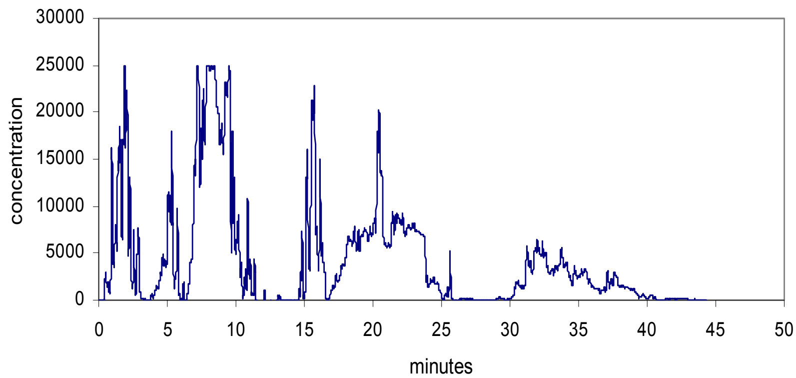
- Cartridges are collected.
- Cartridge barcode is scanned.
- The parameters and test information are included with the analysis results.
- No hand entry.

Trace Gas Analyzer (TGA)

- Continuous analyzer

Does the Sample Created By Filling the Bag Using Pulses Adequately Represent the Average Concentration of the Sampled Tracer?

TGA Plume Plot



Experimental Setup

- PVC pipe suspended horizontally with tubing ports.
- Tracer (SF_6) and ultra pure air injected into one end.
- Flows set with mass flow controllers.
- TGA and PIGS attached to sampling ports.

Three Tests Were Done

1. Test A: "Worst case" scenario.
2. Test B: "Best case" scenario.
3. Test C: "Random pulse" scenario.

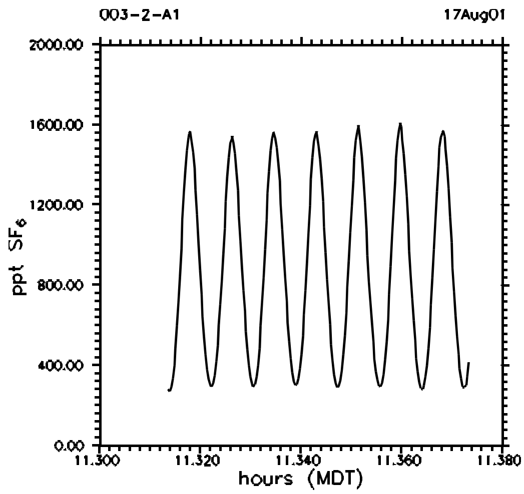
Test A- Worst Case Scenario

- Sampler pulses and SF_6 puffs were the same frequency.
- Samplers in-phase measure higher concentration than out-of-phase.
- Worse than real world because variations were continuous.
- Samplers ran for 50 minutes/bag, pulsing at 30 second intervals.

TGA Monitoring of Tracer

- Two TGAs sampled the gas in the sampling pipe to verify that the concentration puffs were not mixing into a uniform concentration.

TGA Output



Test B- Best Case Scenario

- Same as Test A except sampled almost continuously.
- The difference between Test A and Test B = worst-case effect?

Test C- Random Pulse Scenario

- SF_6 pulsed for 1 second; time between pulses varying randomly.
- Samples pulsed their pumps every 5 seconds.

Results for Each Test

Test	Average Concentration (pptv)	Standard Deviation (pptv)	RSD (%)
A Worst Case	979	155	16
B Best Case	1032	63	6
C Random Pulse	1422	96	7

Worst Case Scenario Produced Greater Imprecision

- Worst Case resulted in the highest RSD of 16%.
- Best Case and Random Pulse had fairly equivalent RSDs of 6% and 7% respectively.

Comparison Between Tests

Test	RPD of Average Concentration (%)	RPD of RSD (%)
Worst Case and Best Cast	8.5	67
Worst Case and Random Pulse	NA	57
Best Case and Random Pulse	NA	11

Random Pulse Mimicked Best Case Very Well

- The RSDs of Best Case and Random Pulse were almost the same.
- The RPD of the RSD between Best Case and Random Pulse was less than the control limit of ± 20%.

Summary

- Not an exhaustive treatment.
- Worst Case 95% confidence intervals were ±32%, only 12% beyond control limits.
- Random Pulse showed no problems.
- Not likely that real world scenarios would cause significant problems.

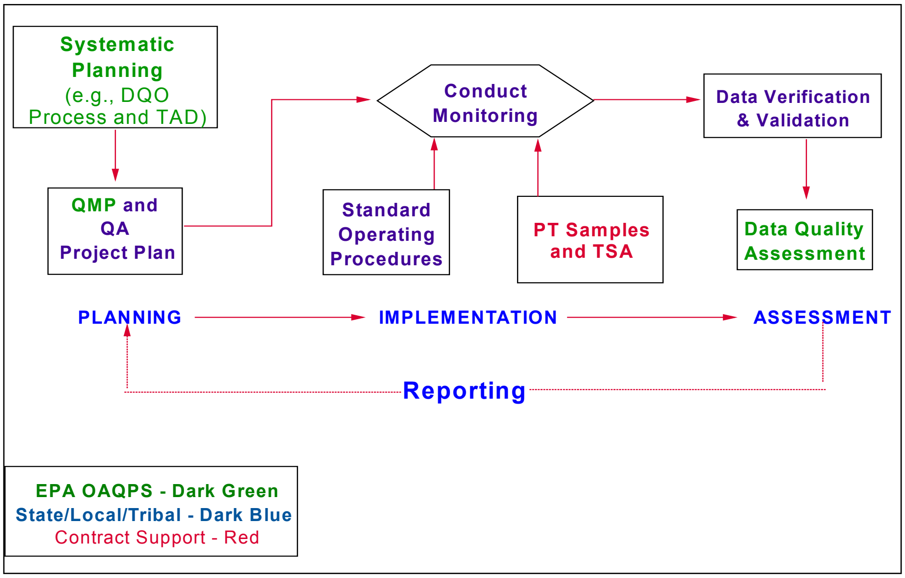
National Air Toxics Trends Stations Quality Assurance System

Dennis Mikel
EPA-OAQPS-EMAD

Outline

- NATTS QA system
  - What is it?
  - Where are we?
  - Issues!
- Results from Proficiency Tests 2001-2003
- Summary

NATTS QA System



QA System Planning

- DQOs: developed 2002 by Battelle Inc.
- QMP – under development
- QAPPs – updates due 9/05
- Technical Assistance Document (TAD): developed by Eastern Research Group (Draft) on AMTIC Website

QA System Implementation

- Proficiency Testing (PT): Quarterly spiked samples to all labs to estimate bias (Mantech Inc.).
  - Will include metals, VOCs and carbonyls
  - First samples out soon!
- Technical System Audits (TSAs): 11 TSAs per year (Battelle Inc.)
  - First TSA to be performed in Philadelphia 4/20/04!
- Cylinder Certifications (EPA ORIA Lab –LV):
  - certify one calibration level cylinder to be sent to 18 NATTS Labs

QA Systems Assessment

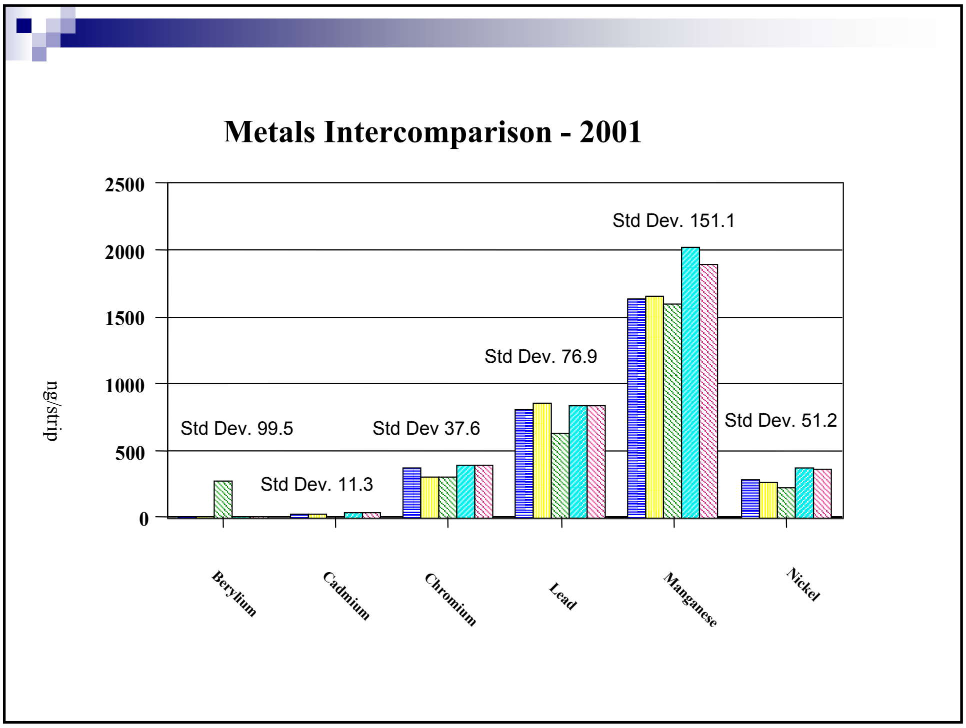
- Annual QA Report: Summarize Precision, Bias, TSA data
- Data Quality Assessments: Summarize whether Precision and Bias goals are met

QA Systems Issues

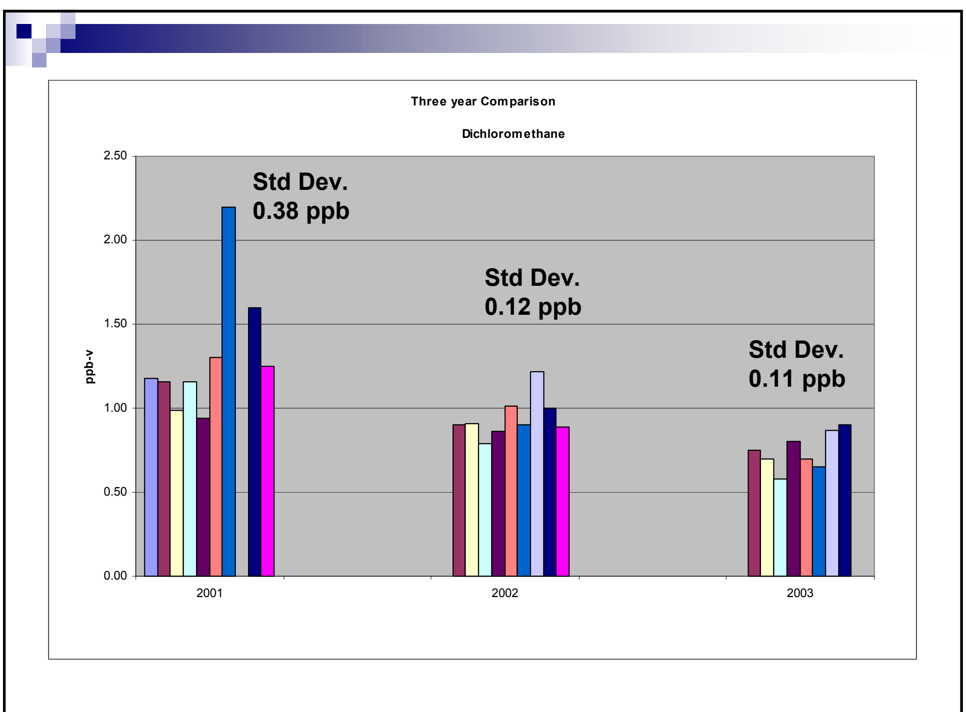
- NIST:
  - Upfront Development
  - Cost
- TAD Development
- Use of Section 105 funding

Proficiency Testing

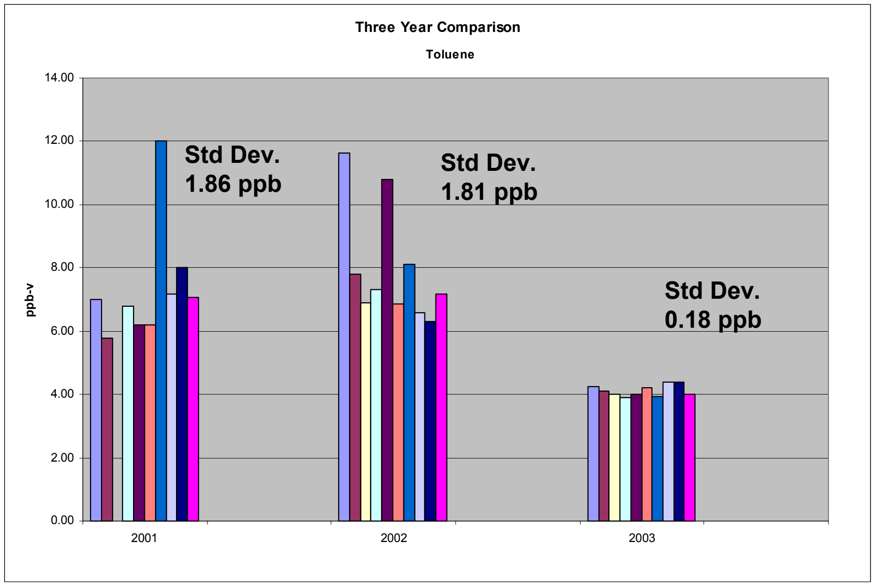
- To date:
  - Pilot study/EPA: 2001 metals
    - Only performed during Pilot study
    - Whole air PM10 samples – 8 x 10 in quartz filter
  - Pilot study/CARB: 2001 to 2003
    - "Whole Air" samples – VOCs
    - Pilot labs continued analysis
  - Region 4 PT results

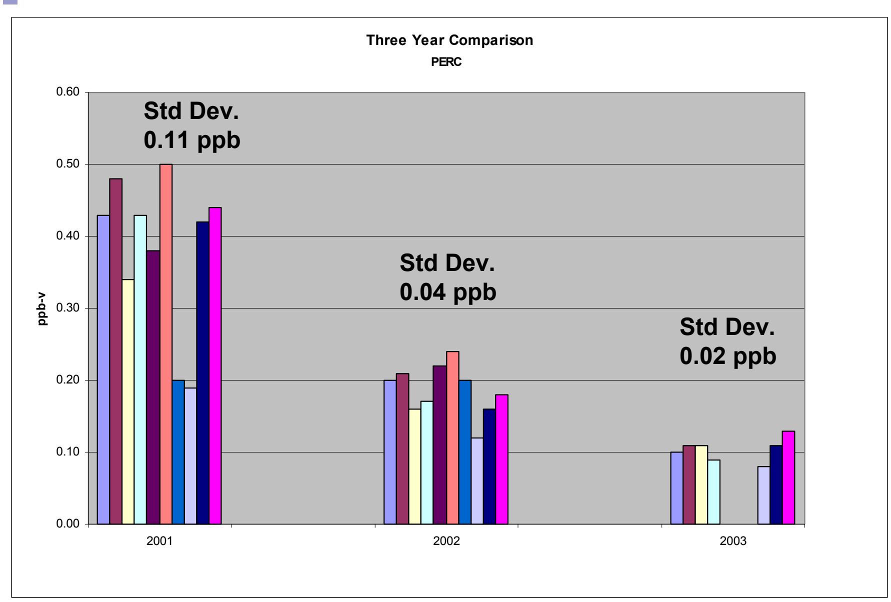


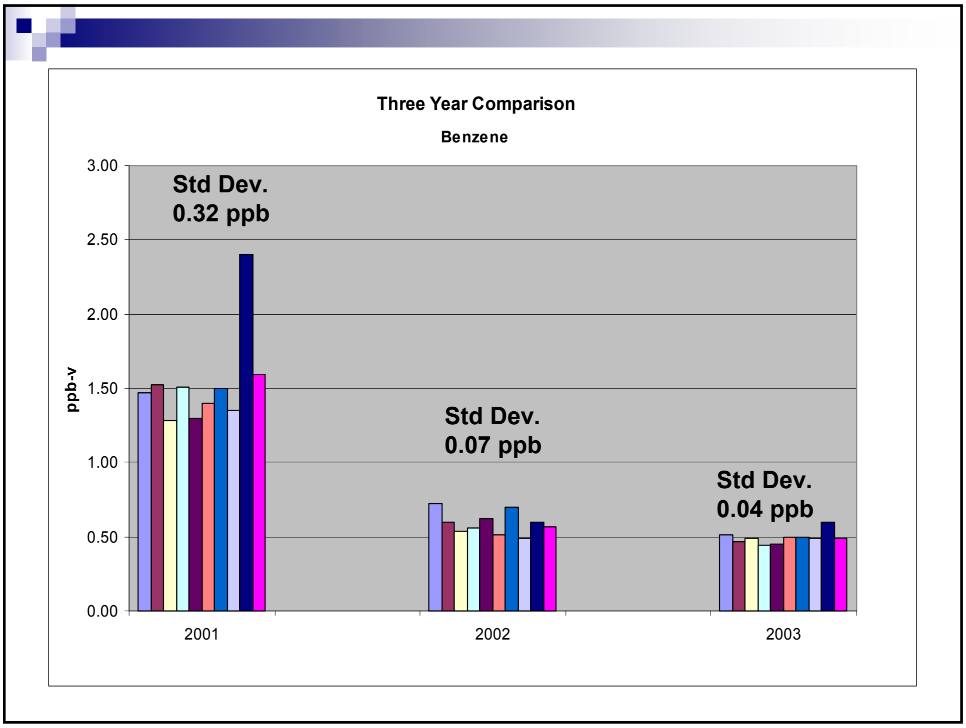
Metals Intercomparison - 2001
ng/strip
2500
2000
1500
1000
500
0
Std Dev. 151.1
Std Dev. 76.9
Std Dev. 99.5
Std Dev 37.6
Std Dev. 51.2
Std Dev. 11.3
Beryllium
Cadmium
Chromium
Lead
Manganese
Nickel

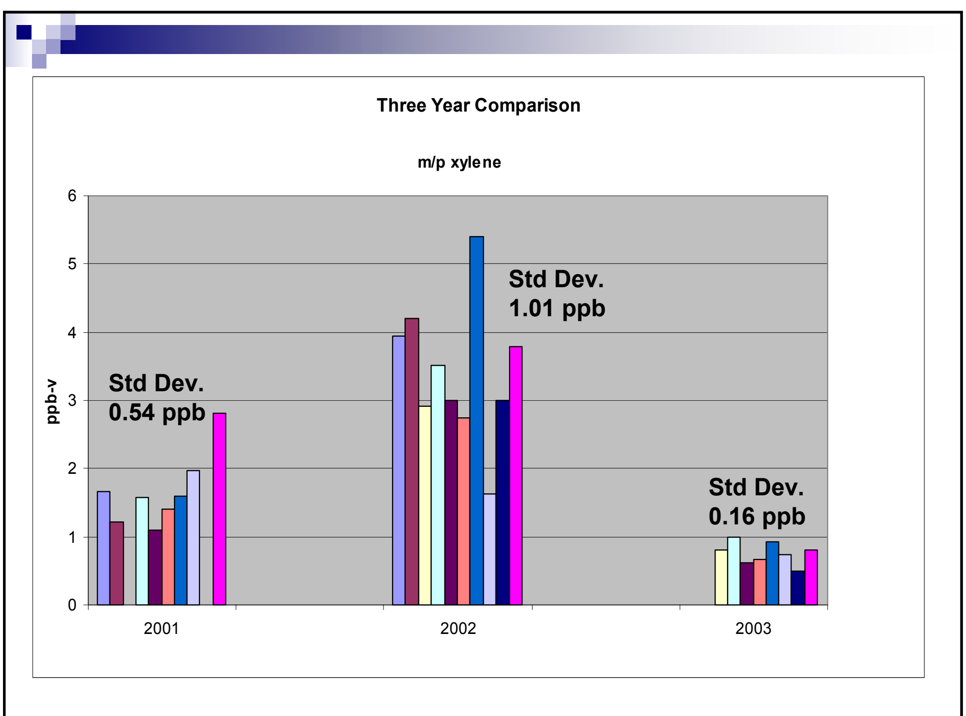


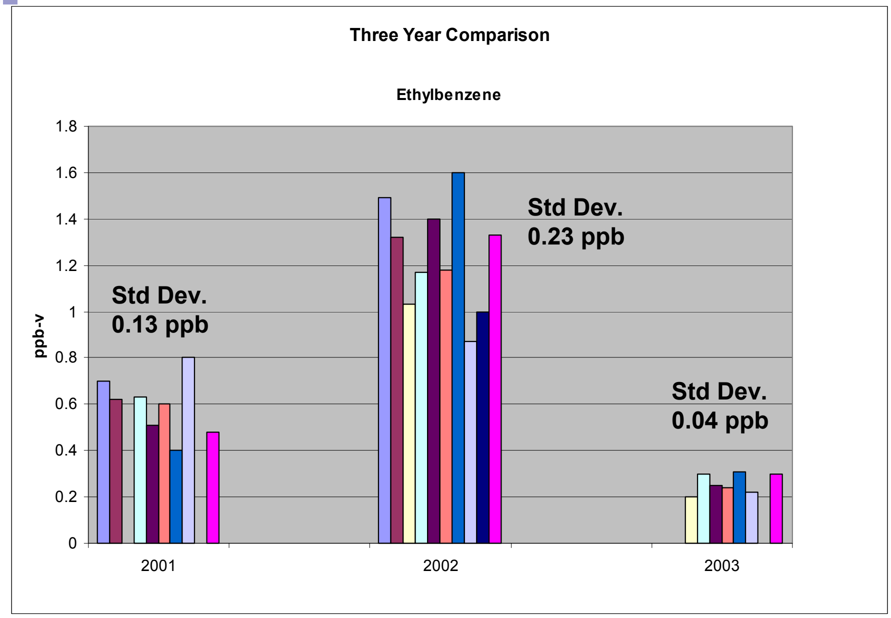
Three year Comparison
Dichloromethane
ppb-v
2.50
2.00
1.50
1.00
0.50
0.00
Std Dev. 0.38 ppb
Std Dev. 0.12 ppb
Std Dev. 0.11 ppb
2001
2002
2003

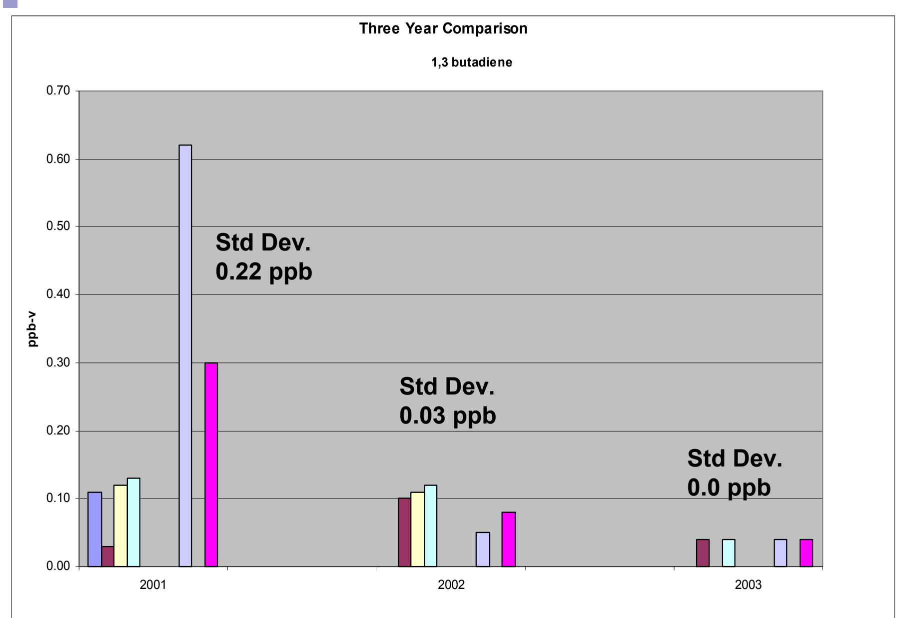

Three Year Comparison
Toluene
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
ppb-v
Std Dev.
1.86 ppb
Std Dev.
1.81 ppb
Std Dev.
0.18 ppb
2001
2002
2003


Three Year Comparison
PERC
0.60
0.50
0.40
0.30
0.20
0.10
0.00
ppb-v
Std Dev.
0.11 ppb
Std Dev.
0.04 ppb
Std Dev.
0.02 ppb
2001
2002
2003


Three Year Comparison
Benzene
Std Dev.
0.32 ppb
Std Dev.
0.07 ppb
Std Dev.
0.04 ppb
ppb-v
3.00
2.50
2.00
1.50
1.00
0.50
0.00
2001
2002
2003


Three Year Comparison
m/p xylene
Std Dev.
1.01 ppb
Std Dev.
0.54 ppb
Std Dev.
0.16 ppb
ppb-v
6
5
4
3
2
1
0
2001
2002
2003


Three Year Comparison
Ethylbenzene
ppb-v
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
2001
2002
2003
Std Dev.
0.13 ppb
Std Dev.
0.23 ppb
Std Dev.
0.04 ppb


Three Year Comparison
1,3 butadiene
ppb-v
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
2001
2002
2003
Std Dev.
0.22 ppb
Std Dev.
0.03 ppb
Std Dev.
0.0 ppb

Summary

- AT QA System is underway!
- On-going PT and TSA
- Whole Air VOCs
  - Extremely useful tool
  - Illustrates problems with “whole air samples”
  - Need referee lab

National Air Toxics Trends Sites

Interim Round-Robin Standard Study

(Very preliminary results)

With multiple labs conducting analysis of the NATTS samples, OAQPS identified the need for PT samples and NIST traceable standard reference materials to provide consistent, accurate, unbiased data.

But getting these things together takes time

As an interim measure OAQPS asked the Region 4 Science and Ecosystem Support Division to prepare a series of standards for analysis by the various NATTS labs.

The EPA Region 4 Science and Ecosystem Support Division prepared VOC standards and purchased trace metal filter standards.

VOC standards were prepared by:

Sallie Hale
(706) 355-8815
&
Jose Rios
(706) 355-8843

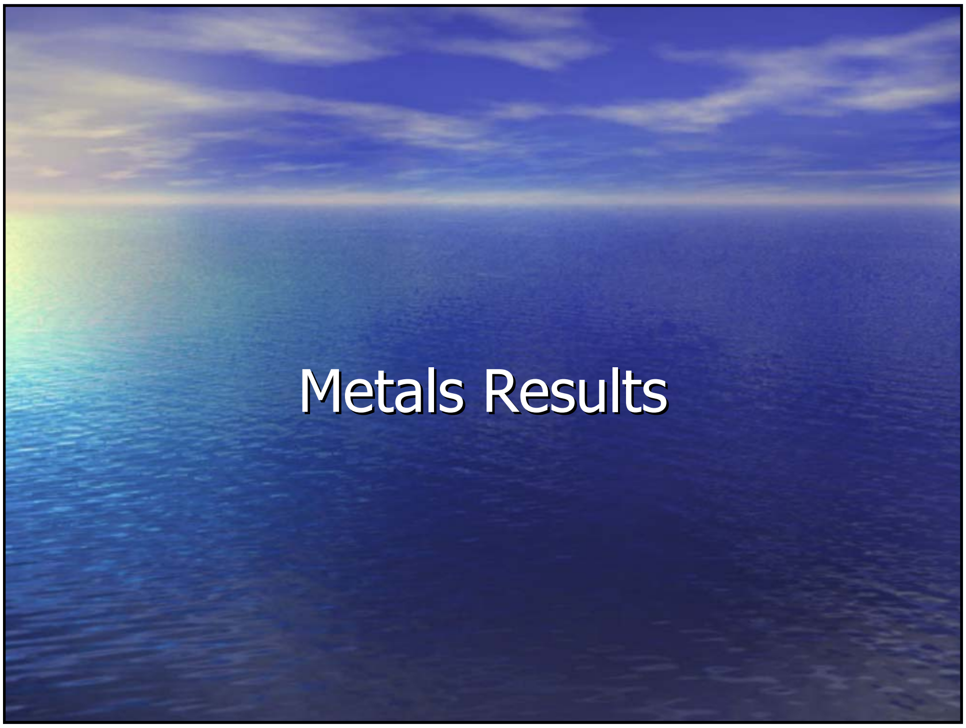
Metals filters were analyzed at SESD by:
Mike Wasko
(706)355-8821

The Trace Metals filter media

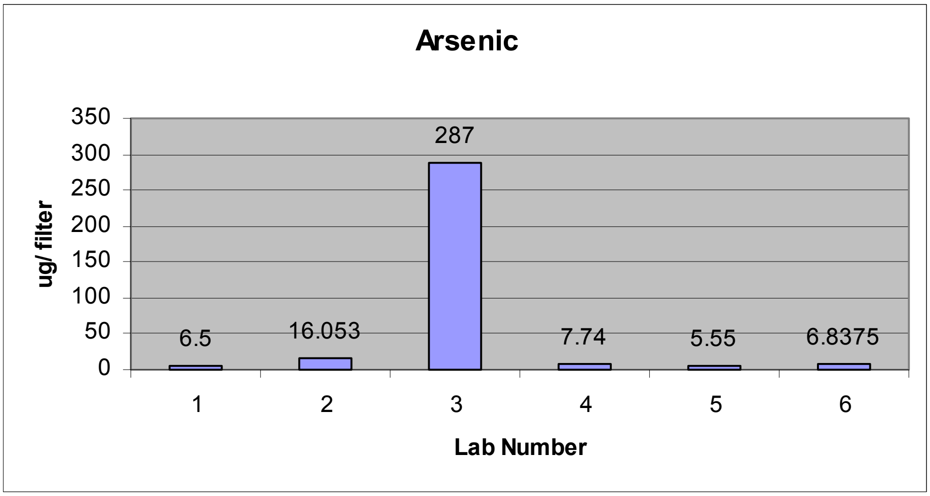
- Certified Reference material purchased from High-Purity Standards designed to meet the requirements of Method 7300.
- Cellulose ester type filter spike with a component standard solution containing 16 elements.
- Blank filters were also purchased.

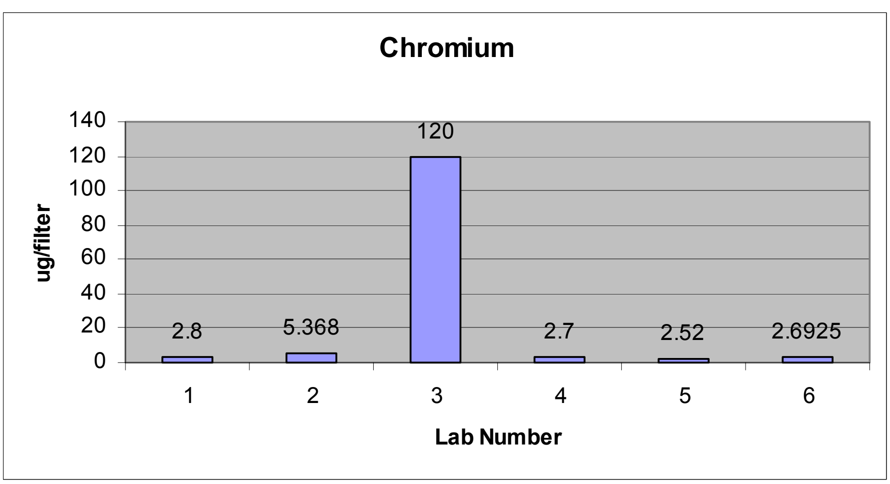
VOC Standards

- Spectra Gas nominal 100 ppbv standard
- Dilute the standard in humidified, high purity nitrogen
- Equipment used:
  - Entech Model 4600 Dynamic Diluter
  - Entech Model 7100 Preconcentrator
  - Entech Model 7016CA Autosampler
  - Agilent 6890 Gas Chromatograph
  - Agilent 5973N Mass Spectrometer

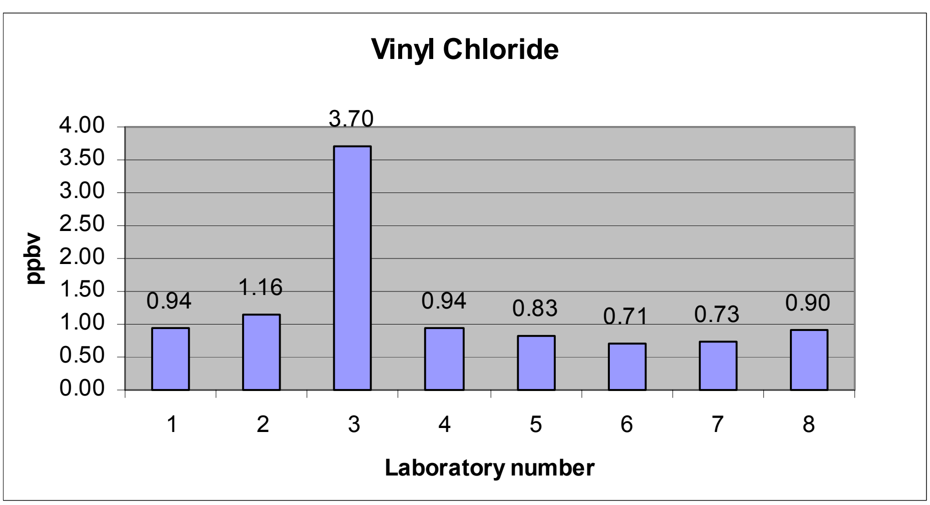


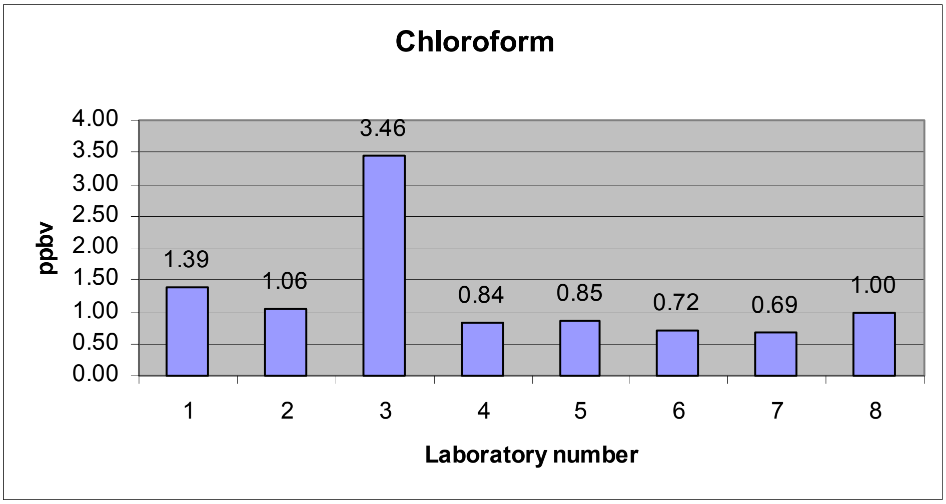
Lab #	1	2	3	4	5	6
Element	Proficiency Sample ug/filter	Proficiency Sample ug/filter	Proficiency Sample ug/filter	Proficiency Sample ug/filter	Proficiency Sample ug/filter	Proficiency Sample ug/filter
Arsenic	6.5	16.053	287	7.74	5.55	6.8375
Barium	2.4				2.63	
Beryllium	0.95	1.584	45	1.03	0.89	1.01
Cadmium	0.9	2.144	49	1.01	0.87	1.0275
Cobalt	2.4	5.942	120		2.39	
Copper	2.6				2.68	
Iron					3.72	
Manganes	1	2.659	46	1.01	1.16	0.98
Lead	2.6	4.896	124	2.44	2.42	2.7375
Antimony		<0.005				
Selenium	2.4	5.787	113			
Chromium	2.8	5.368	120	2.7	2.52	2.6925
Nickel	2.7	4.861	129	2.6	2.43	2.1925
Vanadium	2.6				2.42	
Zinc	3.4		266		12.49	


Arsenic
350
300
250
200
150
100
50
0
ug/ filter
287
6.5
16.053
7.74
5.55
6.8375
1
2
3
4
5
6
Lab Number

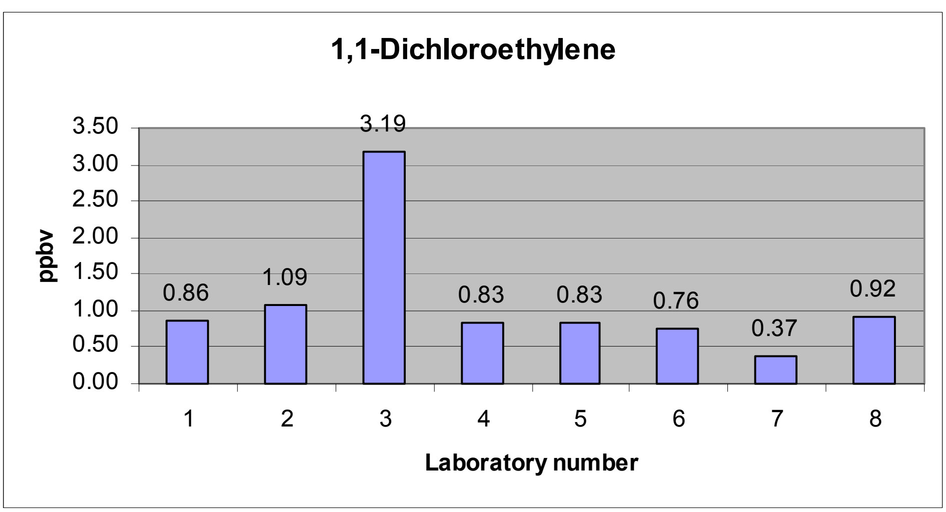

Chromium
140
120
100
80
60
40
20
0
ug/filter
120
2.8
5.368
2.7
2.52
2.6925
1
2
3
4
5
6
Lab Number

VOC Results

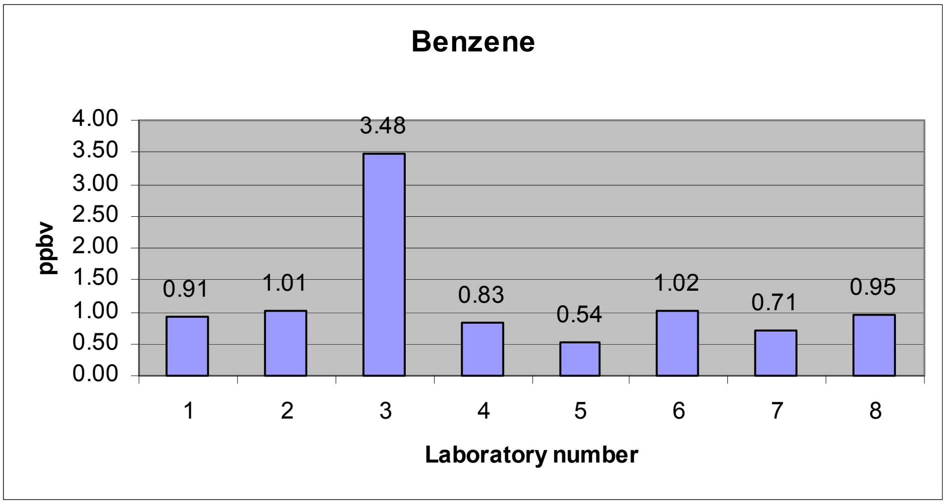

Vinyl Chloride
ppbv
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
0.94
1.16
3.70
0.94
0.83
0.71
0.73
0.90
1
2
3
4
5
6
7
8
Laboratory number



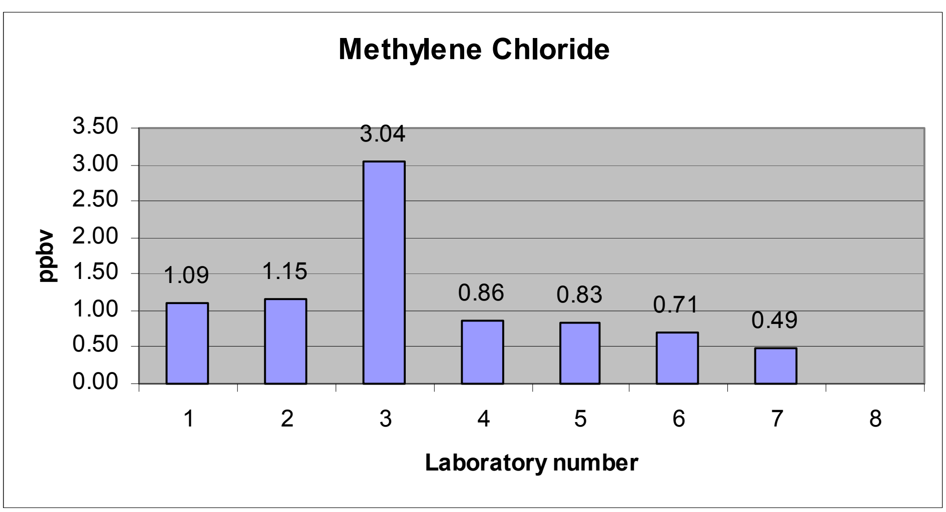
Chloroform
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
ppbv
1.39
1.06
3.46
0.84
0.85
0.72
0.69
1.00
1
2
3
4
5
6
7
8
Laboratory number



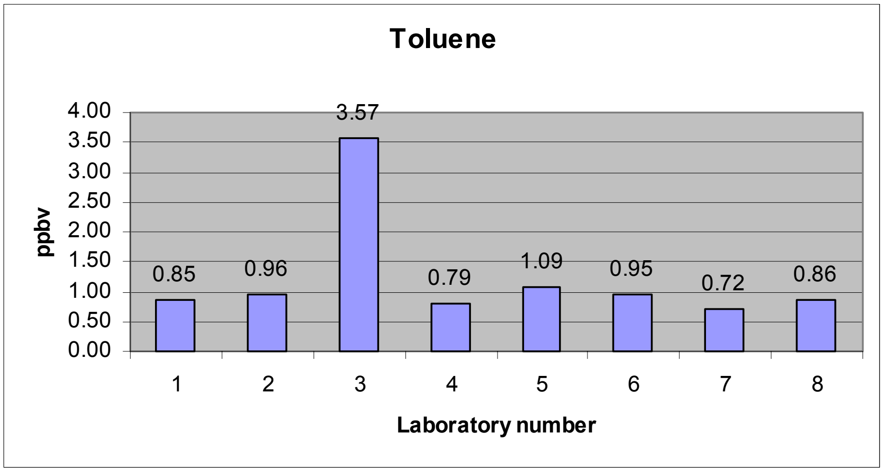
1,1-Dichloroethylene
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
ppbv
0.86
1.09
3.19
0.83
0.83
0.76
0.37
0.92
1
2
3
4
5
6
7
8
Laboratory number

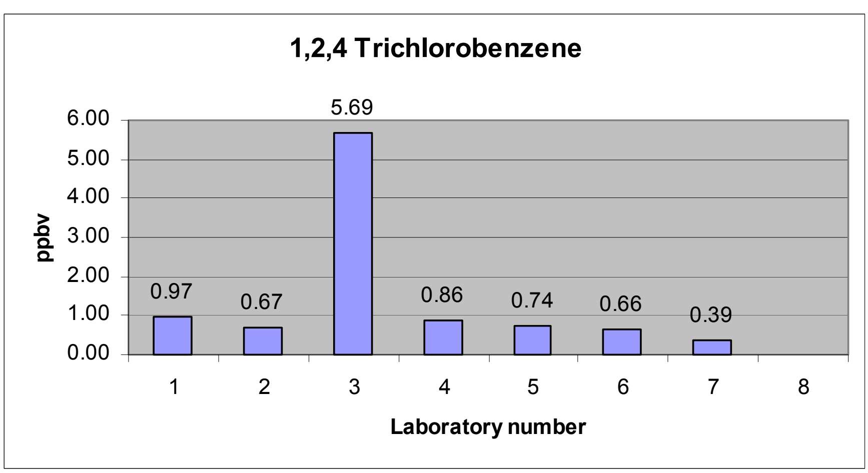


Benzene
ppbv
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
0.91
1.01
3.48
0.83
0.54
1.02
0.71
0.95
1
2
3
4
5
6
7
8
Laboratory number



Methylene Chloride
ppbv
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
1.09
1.15
3.04
0.86
0.83
0.71
0.49
1
2
3
4
5
6
7
8
Laboratory number


Toluene
ppbv
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
0.85
0.96
3.57
0.79
1.09
0.95
0.72
0.86
1
2
3
4
5
6
7
8
Laboratory number


1,2,4 Trichlorobenzene
ppbv
6.00
5.00
4.00
3.00
2.00
1.00
0.00
0.97
0.67
5.69
0.86
0.74
0.66
0.39
1
2
3
4
5
6
7
8
Laboratory number

Conclusions

- PT samples and NIST traceable standards are essential for producing high quality, consistent data.
- May need PT samples available several times per year. The National Environmental Laboratory Accreditation Conference (NELAC) recommends twice per year.
- When sending PT samples, must be VERY clear with directions on how to handle calculations and reporting..

Tribal Air Monitoring Support Center

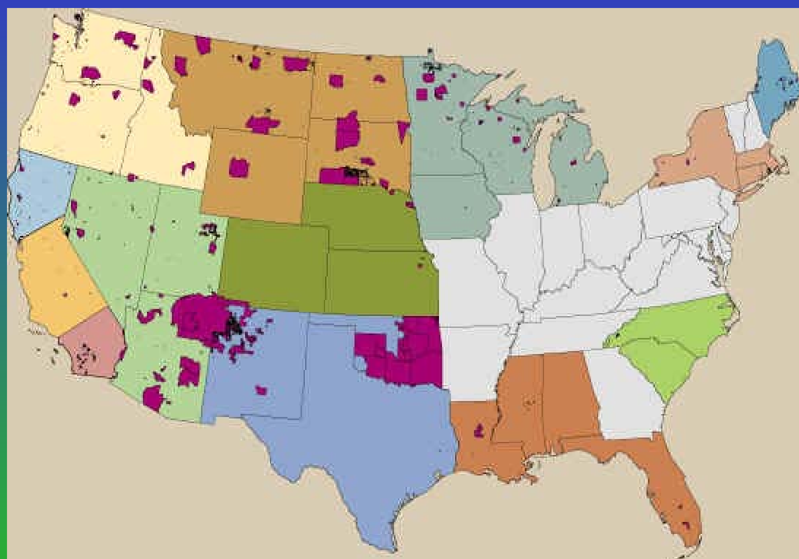
Contributions of Tribal Environmental Agencies to Quality Assurance in Environmental Monitoring

Melinda Ronca-Battista

TAMS Center:

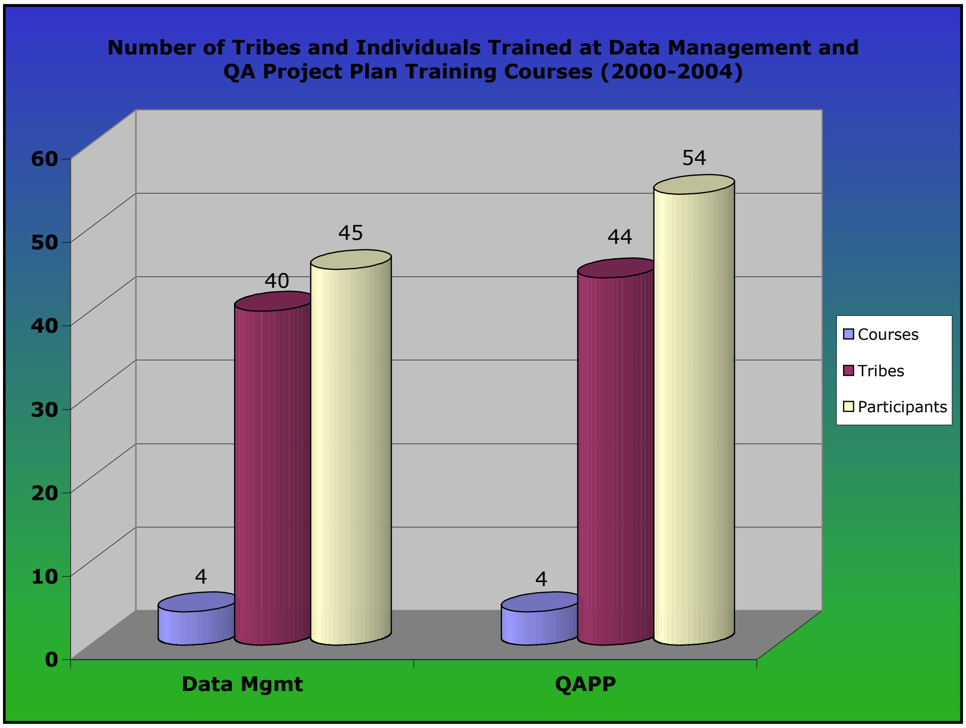
- Northern Arizona University
- US EPA
- Mission is to develop tribal capacity to assess, understand and prevent environmental impacts that adversely affect health, cultural, and natural resources

Tribal lands in continental US:

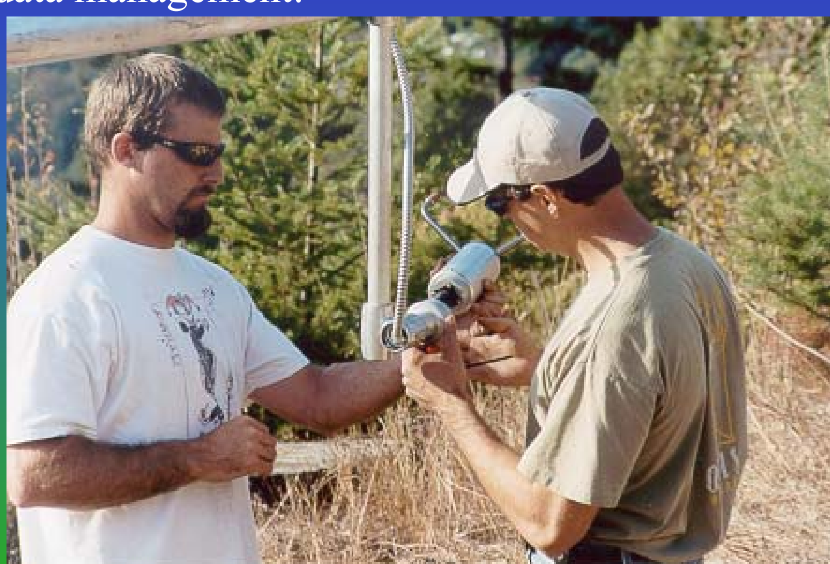


TAMS Learning Center in Las Vegas:

Hands-on learning techniques:



On-site field assistance with instrumentation and data management:



What does "Contributions to QA in Environmental Monitoring" mean"?

- QA is the system "to obtain data of the type and quality needed by the client"
- Positive contribution is made when that system is efficient and flexible
- Environmental monitoring is improved when the organization eats, drinks, and breathes the environment

Tribal contributions to QA:

- Two major categories of characteristics that differentiate tribal environmental agencies from other "reporting organizations"
  - ***Small and relatively young***
  - ***Philosophically different than local and state***

How small organizations contribute to QA:

- Efficient:
  - QA documents are written and understood by implementers
  - Improvement cycle fast
- Flexible:
  - Community concerns can be quickly addressed
  - Special projects respond to needs
  - Updated information from vendors and US EPA can be quickly incorporated in program

Young organizations:

- Are starting fresh, with no entrenched procedures
- Have benefit of recent advances and simplifications in guidelines and instrumentation
- Have benefit of learning from others

Philosophical differences between Tribes and other reporting organizations

1. Sovereign entities, with status parallel to that of a nation
   - Can write their own environmental laws or adopt federal or state laws
   - Primary reporting is to Tribal Council and community, rather than US EPA or state
   - Not limited by existing environmental protection structure

(disadvantages of not being subject to existing federal, state, or local laws):

Philosophical differences, cont.:

2. Cultural value of environment:

– mountains, geologic formations, water bodies, vistas form the definitions of the people
– Ceremonies use and refer to environment
– Various clans from different regions are associated with the environment in that region
– Being "of the earth" is taken literally, as in "of *this* earth"

Special project to identify land contaminated with waste from uranium ore trucking:

Philosophical differences, cont.

3. Subsistence living

- fish, in oceans, lakes, and rivers
- crops, such as corn, wild rice
- hunting
- livestock

Hopi dry corn farming

Philosophical differences, final

4. Generational context of environmental protection:
   - Respect for ancestors who are of the land
   - Care of future generations part of explicitly stated philosophies of tribal environmental agencies
   - Identification with land is permanent

***Yisk'aaz* is the Navajo word for quality. It is a word that evokes a system whose parts fit together, which work in unison, and are in balance. To walk in beauty is the Navajo idea of a life in balance. Quality assurance in the Navajo view of the world is an integral part of the system. All parts of the system need work perfectly for *Yisk'aaz* (quality) to be achieved.**

Stanley Edison, Navajo Nation Superfund Program QAPP

Coarse PM Methods Evaluation Study

Study Design and Results

R. Vanderpool, T. Ellestad, P. Solomon, and M. Harmon
US EPA – ORD – NERL

S. Natarajan, C. Noble, and R. Murdoch
Research Triangle Institute

J. Ambs (Rupprecht & Patashnick Co., Inc.),
J. Tisch (Tisch Environmental, Inc.), and G. Sem (TSI Inc.)

Background

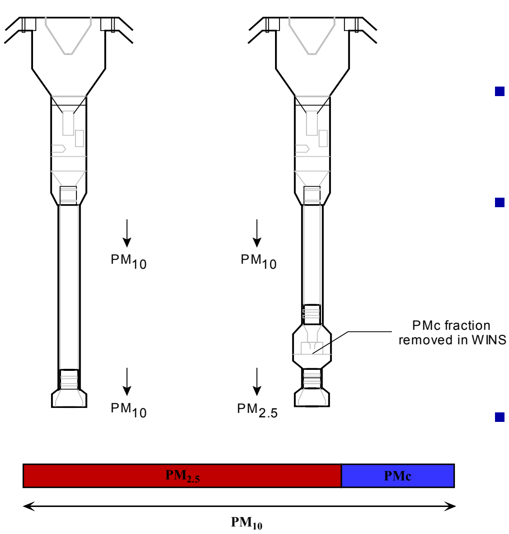
- Since the 1997 $PM_{2.5}$ promulgation, the U.S. courts have reviewed subsequent litigation and ruled that the PM_{10} metric is a "poorly matched indicator" because it includes the $PM_{2.5}$ fraction. EPA has consented to establish separate air quality standards for the fine and coarse fractions of PM_{10}

- PMc is inherently more difficult to accurately measure than either $PM_{2.5}$ or PM_{10}. Measurement issues (e.g. losses of large particles) may result in less precise PMc measurements than either $PM_{2.5}$ or PM_{10} measurements

Study Objectives

- Evaluate the field performance of leading methods for monitoring the coarse fraction of PM_{10} ($PMc = PM_{10} - PM_{2.5}$)
- Evaluate samplers which are either already commercially available or in their final stages of development
- Include both filter-based (time-integrated) and semi-continuous measurement methods

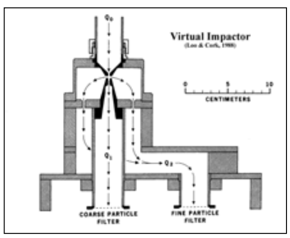
$PM_{2.5}$ and PM_{10} FRM Samplers



$PMc = PM_{10} - PM_{2.5}$

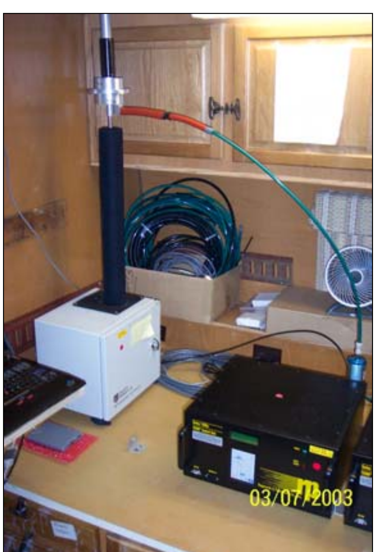
- Standard low-vol PM_{10} inlets aspirating at 16.7 lpm (actual conditions)
- $PM_{2.5}$ aerosol fractionation using a WINS equipped with DOS impaction oil
- Filters were conditioned at 22C and 35% RH, analyzed gravimetrically. Post-sampling filters archived at -30C for subsequent chemical analysis
- 3 FRM pairs from BGI, R&P, and Thermo-Andersen equipped with teflon filters (4th FRM pair equipped with quartz filters)

R&P Partisol-Plus 2025 Dichot



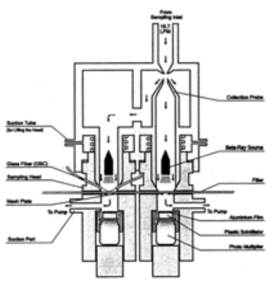
- Standard PM_{10} inlet aspirating at 16.7 lpm (actual)
- Aerosol fractionation by custom virtual impactor (15 lpm and 1.67 lpm)
- $PM_{2.5}$ and PMc mass collected on 47 mm teflon filters for gravimetric analysis
- Sequential sampler with multi-day capability
- 4 units used in our study (3 teflon and 1 quartz)

R&P Coarse Particle TEOM



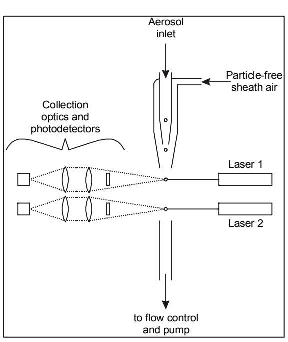
- Modified PM_{10} inlet aspirating at 50 lpm (actual)
- PM_{10} aerosol is fractionated by a custom virtual impactor (2 lpm coarse flow and 48 lpm fine flow)
- PMc fraction is heated to 50 C to remove particle bound water
- Coarse aerosol is collected and quantified by a standard TEOM sensor
- 3 units used in our study

Tisch SPM-613D Dichot Beta Gauge



- Standard PM_{10} inlet aspirating at 16.7 lpm (~std)
- Aerosol heated if <25C
- Aerosol fractionation by custom virtual impactor
- $PM_{2.5}$ and PMc mass collected on polyflon tape roll
- $PM_{2.5}$ and PMc mass quantified hourly using separate beta sources and detectors
- 3 units used in our study

TSI Model 3321 Aerodynamic Particle Sizer



- Standard PM_{10} inlet aspirating at 16.7 lpm (actual)
- Isokinetic fraction of PM_{10} aerosol removed at 5 lpm and enters the APS inlet
- APS sizes individual particles aerodynamically using time of flight approach
- Single particle volume converted to mass using mean density provided by user
- Total aerosol mass is sum of individual particle masses
- APS provides only PMc; not applicable for $PM_{2.5}$ or PM_{10}
- Only sampler in study which provides detailed size distribution information
- 2 units used in our study

Mobile Sampling Platform (Side View)

Study Design

- Using 22 hour, daily sampling periods for comparisons (11 am to 9 am local time)
- Chemical analysis (XRF, IC, thermal optical) of selected archived filters will provide particle composition, which may help explain observed sampler performance
- The difference method will be used as the basis of comparison for the study

Sampler Performance Issues

- Relative bias compared to collocated FRMs
- Precision (2 or 3 samplers of each type)
- Field reliability
- Evaluation under a wide range of weather conditions and aerosol types

QA/QC Initiatives

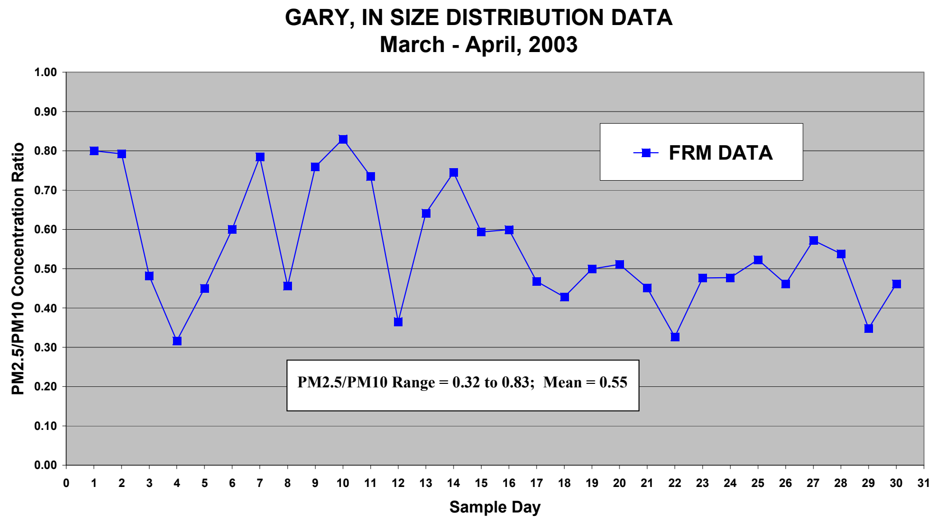
- Sampler manufacturers were very involved in the study and allowed to verify the working condition of their respective samplers at each sampling site
- Sampling and fractionation components were cleaned prior to sampling at each site
- NIST-traceable sampler calibration equipment was used for all sampler calibrations and audits
- Three performance audits and three field blank tests were conducted at each site
- Replicate weighings were conducted for all samples
- Weighings were done on-site (before shipping) and at EPA's RTP weighing facility to measure PM losses during shipment

Study Sites

- RTP, NC (10 days of shakedown tests, Jan. 2003)
- Gary, IN (30 days of tests under cold, snow/rain, variable $PM_{2.5}/PM_{10}$ ratios, March-April, 2003)
- Phoenix, AZ (30 days of tests under hot, dusty conditions, consistently low $PM_{2.5}/PM_{10}$ ratios, May-June, 2003)
- Riverside, CA (30 days of tests under warm conditions, higher $PM_{2.5}/PM_{10}$ ratios than Phoenix, July-August, 2003)
- Phoenix, AZ (15 days of follow-up tests, January 2004)

Gary, IN

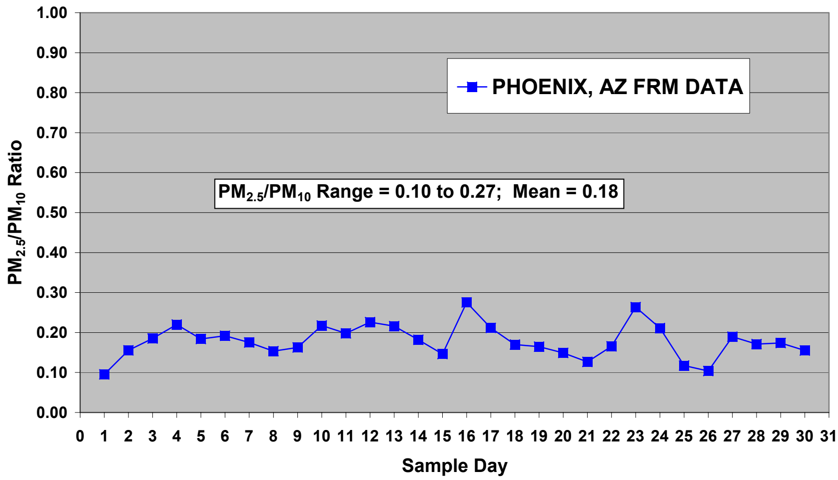
Mean daily temperature = 4.6 C


GARY, IN SIZE DISTRIBUTION DATA
March - April, 2003
PM2.5/PM10 Concentration Ratio
Sample Day
FRM DATA
PM2.5/PM10 Range = 0.32 to 0.83; Mean = 0.55

Phoenix, AZ

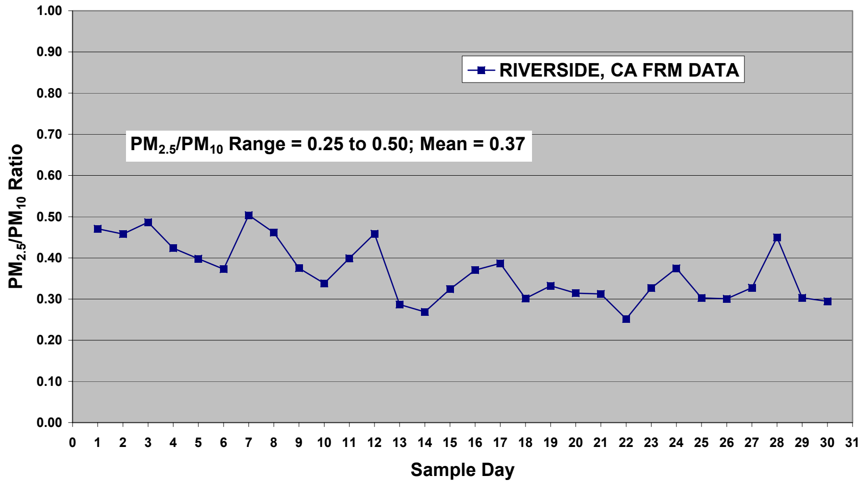

Mean daily temperature = 32.3 C
Phoenix

PHOENIX, AZ SIZE DISTRIBUTION DATA
May - June, 2003

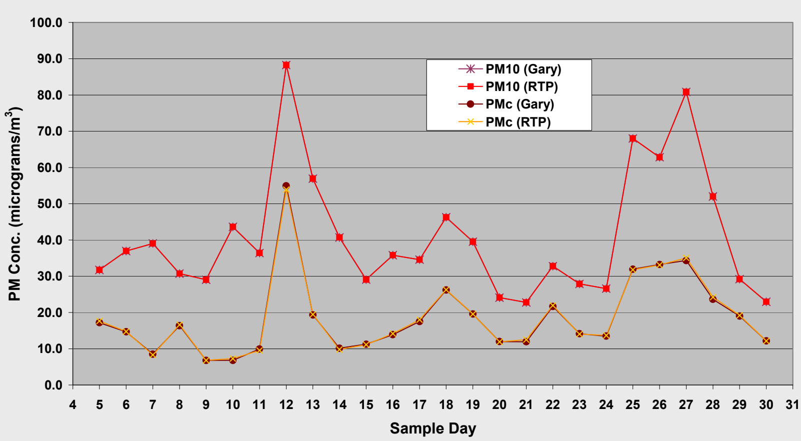


Riverside, CA
UCR Ag Ops Facility

RIVERSIDE, CA SIZE DISTRIBUTION
July - August, 2003

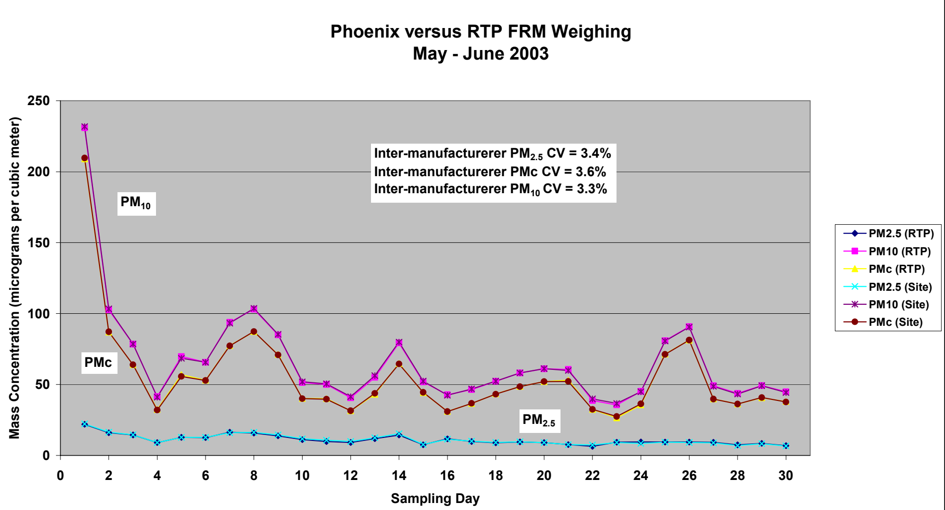


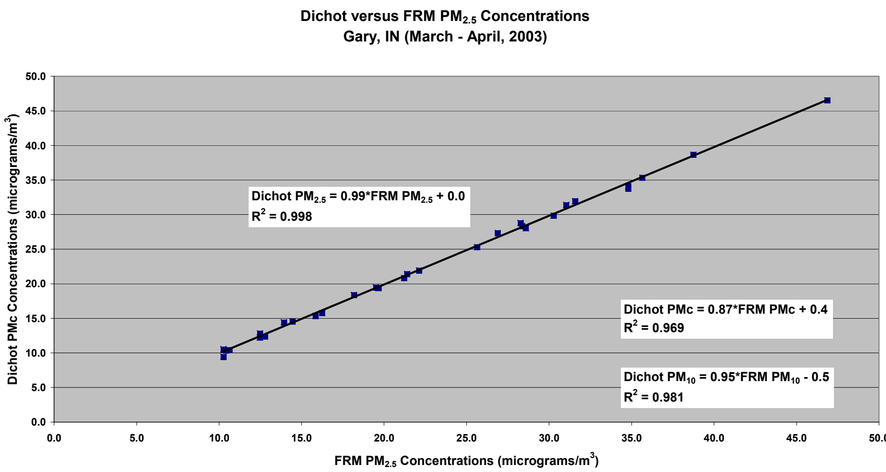
FRM MEASUREMENTS - GARY vs RTP WEIGHING
Gary, IN (March - April, 2003)



RTP/Site Test Results for Integrated Samplers

		SAMPLING SITE				
Metric	Sampler	RTP/Gary	RTP/Phoenix (May–June, 2003)	RTP/Riverside	RTP/Phoenix (Jan. 2004)	Mean
$PM_{2.5}$	FRM	0.97	0.99	1.00	1.00	0.99
	R&P Dichot (seq)	0.95	1.00	1.00	1.00	0.99
	R&P Dichot (man)				1.00	1.00
	Andersen Dichot				0.99	0.99
PM_{10}	FRM	0.98	1.00	1.00	1.00	1.00
	R&P Dichot (seq)	0.97	1.00	0.98	1.01	0.99
	R&P Dichot (man)				1.01	1.01
	Andersen Dichot				0.98	0.98
PMc	FRM	1.00	1.00	1.00	1.00	1.00
	R&P Dichot (seq)	0.99	1.00	0.96	1.02	0.99
	R&P Dichot (man)				1.02	1.02
	Andersen Dichot				0.97	0.97
					Mean =	0.99

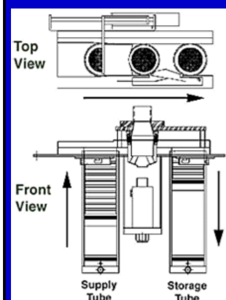




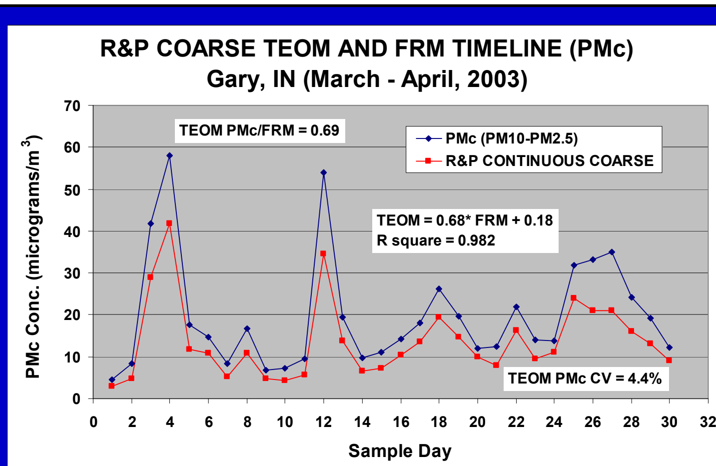
R&P Dichots vs. FRM

Metric	Gary, IN	Phoenix, AZ	Riverside, CA
$PM_{2.5}$	Slope = 0.99 Int. = +0.0 R^2 = 0.998 Ratio to FRM = 0.99	Slope = 1.24 Int. = -1.6 R^2 = 0.97 Ratio to FRM = 1.09	Slope = 0.998 Int. = +0.0 R^2 = 0.995 Ratio to FRM = 1.00
PMc	Slope = 0.87 Int. = +0.39 R^2 = 0.969 Ratio to FRM = 0.89	Slope = 0.70 Int. = +5.0 R^2 = 0.98 Ratio to FRM = 0.79	Slope = 0.95 Int. = +0.25 R^2 = 0.98 Ratio to FRM = 0.96
PM_{10}	Slope = 0.95 Int. = -0.47 R^2 = 0.981 Ratio to FRM = 0.94	Slope = 0.75 Int. = +5.9 R^2 = 0.98 Ratio to FRM = 0.84	Slope = 1.00 Int. = -1.21 R^2 = 0.99 Ratio to FRM = 0.97

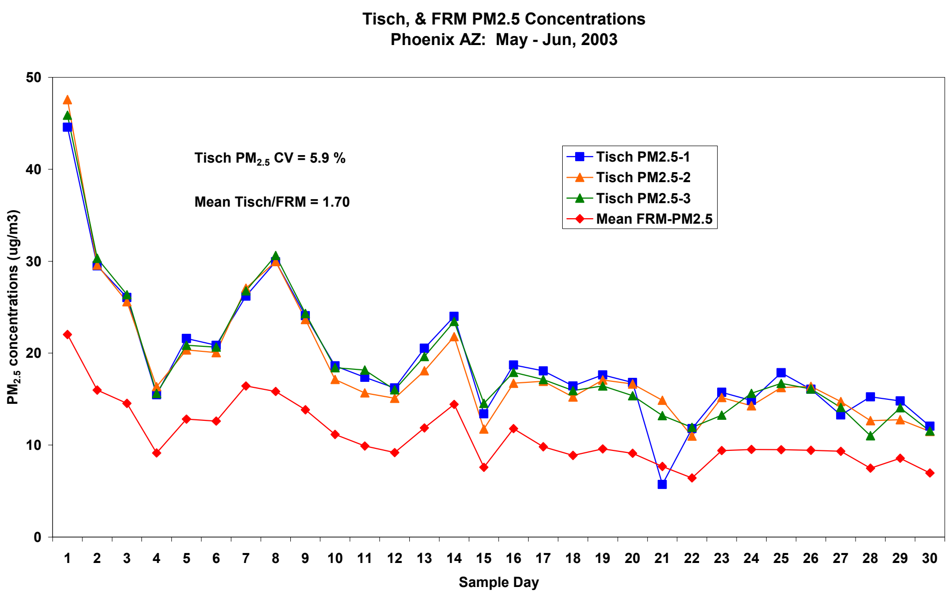
SEQUENTIAL VS. MANUAL DICHOTS
PHOENIX, AZ (JAN 2004)



Metric	Sequential Dichot	Manual Dichot
$PM_{2.5}$	Slope = 1.09 Int. = -0.32 R^2 = 0.982 Ratio to FRM = 1.07	Slope = 1.03 Int. = +0.10 R^2 = 0.982 Ratio to FRM = 1.04
PMc	Slope = 0.84 Int. = +1.5 R^2 = 0.971 Ratio to FRM = 0.90	Slope = 1.02 Int. = -0.08 R^2 = 0.996 Ratio to FRM = 1.00
PM_{10}	Slope = 0.89 Int. = +1.9 R^2 = 0.976 Ratio to FRM = 0.94	Slope = 1.03 Int. = -0.50 R^2 = 0.997 Ratio to FRM = 1.01

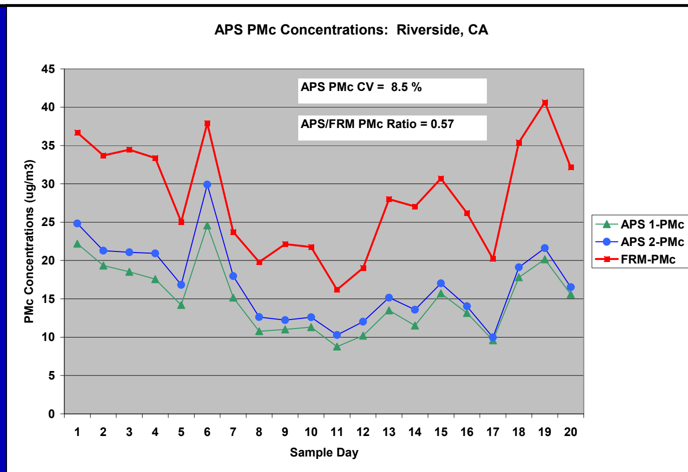


Metric	Gary, IN	Phoenix, AZ (May – June, 2003)	Riverside, CA	Phoenix, AZ (Jan 2004)
PMc	Slope = 0.68 Int. = +0.18 R^2 = 0.982 CV = 4.4% Ratio to FRM = 0.69	Slope = 0.79 Int. = +12.8 R^2 = 0.951 CV = 6.6% Ratio to FRM = 1.05	Slope = 0.74 Int. = -0.64 R^2 = 0.948 CV = 1.7% Ratio to FRM = 0.76	Slope = 0.77 Int. = +0.70 R^2 = 0.995 CV = 2.6% Ratio to FRM = 0.80



Tisch Beta Gauge Dichot vs the FRM

Metric	Gary, IN	Phoenix, AZ (May – June, 2003)	Riverside, CA	Phoenix, AZ (Jan 2004)
$PM_{2.5}$	Slope = 1.17 Int. = +1.6 R^2 = 0.945 Ratio to FRM = 1.26	Slope = 2.03 Int. = -3.4 R^2 = 0.946 Ratio to FRM = 1.70	Slope = 2.07 Int. = -6.9 R^2 = 0.904 Ratio to FRM = 1.64	Slope = 1.43 Int. = -0.11 R^2 = 0.939 Ratio to FRM = 1.43
PMc	Slope = 0.885 Int. = +0.34 R^2 = 0.978 Ratio to FRM = 0.91	Slope = 0.92 Int. = +5.9 R^2 = 0.995 Ratio to FRM = 1.04	Slope = 1.17 Int. = -2.7 R^2 = 0.957 Ratio to FRM = 1.08	Slope = 0.99 Int. = +1.66 R^2 = 0.994 Ratio to FRM = 1.05
PM_{10}	Slope = 1.02 Int. = +2.5 R^2 = 0.987 Ratio to FRM = 1.09	Slope = 1.02 Int. = +7.8 R^2 = 0.996 Ratio to FRM = 1.16	Slope = 1.53 Int. = -10.6 R^2 = 0.880 Ratio to FRM = 1.29	Slope = 1.07 Int. = +2.9 R^2 = 0.998 Ratio to FRM = 1.14



Metric	Gary, IN	Phoenix, AZ (May – June, 2003)	Riverside, CA	Phoenix, AZ (Jan 2004)
PMc	Slope = 0.42 Int. = +0.48 R^2 = 0.80 Ratio to FRM = 0.42	Slope = 0.56 Int. = -0.20 R^2 = 0.99 Ratio to FRM = 0.55	Slope = 0.66 Int. = -2.3 R^2 = 0.82 Ratio to FRM = 0.58	Slope = 0.61 Int. = +0.16 R^2 = 0.993 Ratio to FRM = 0.62

Summary of Results

(independent of site)

- FRMs show strong inter-manufacturer precision (CV<6% for all three metrics) with no tendency for producing negative PMc values
- Filter-based dichots show strong precision (CV<5% for all metrics)
- Site weighing results agree closely with RTP results
- Precision of the semi-continuous samplers ranged from very good to acceptable
- Correlation (as R^2) of semi-continuous samplers with the collocated FRMs is usually strong (>0.95)

R&P 2025 Dichot

- Dichots showed strong inter-sampler precision (CV<4%) at all sites for all three PM metrics.
- Strong correlation (R^2 > 0.980) was observed between the dichots and the collocated FRMs at all sites.
- With the exception of Phoenix, $PM_{2.5}$ dichot concentrations agreed well with the $PM_{2.5}$ FRMs.
- Except during the Riverside tests, the dichots underestimated PMc concentrations by >10%. Mass balance calculations showed that 16% of the aspirated PM_{10} aerosol in Phoenix was unaccounted for in the 2025 dichot. January 2004 follow-up tests in Phoenix indicated that particle losses can occur during post-sampling movement of the coarse particle cassette in the sequential sampler.

R&P PMc TEOM

- Excellent inter-sampler precision was observed among the R&P PMc TEOMs at all sites (mean CV = 4%).
- Correlation between the PMc TEOMs and the collocated FRMs was strong ($R^2 \geq 0.95$) at all sites.
- With the exception of Phoenix in 2003, the coarse TEOMs produced PMc concentrations 20% to 30% lower than the FRMs. Depending upon aerosol size distribution, the ~ 9.0 to 9.5 micrometer cutpoint of the TEOM's 50 lpm inlet may account for an appreciable fraction of the observed difference.

Tisch Dichotomous Beta Gauge

- Acceptable inter-sampler precision was observed for the Tisch samplers for all three metrics at all three sites.
- Good correlation (R^2 > 0.880 for all metrics) was observed between the Tisch samplers and the collocated FRMs at all sites.
- The Tisch sampler consistently overestimated (25% to 70%) the $PM_{2.5}$ concentration at all sampling sites. The units provided lower bias (<10%) when measuring ambient PMc concentrations, whereas PM_{10} was overestimated by about 10-30%.

TSI Aerodynamic Particle Sizer

- The two APS units showed acceptable precision at all sampling sites and provided detailed size distribution information not provided by the other samplers involved in the study.
- PMc concentrations measured by the APS units typically “tracked” concentrations measured by the collocated FRM samplers.
- Independent of sampling site, the APS units typically underestimated PMc concentrations by a factor of two. This field behavior is consistent with published performance tests conducted in the laboratory under controlled conditions.

Future Work

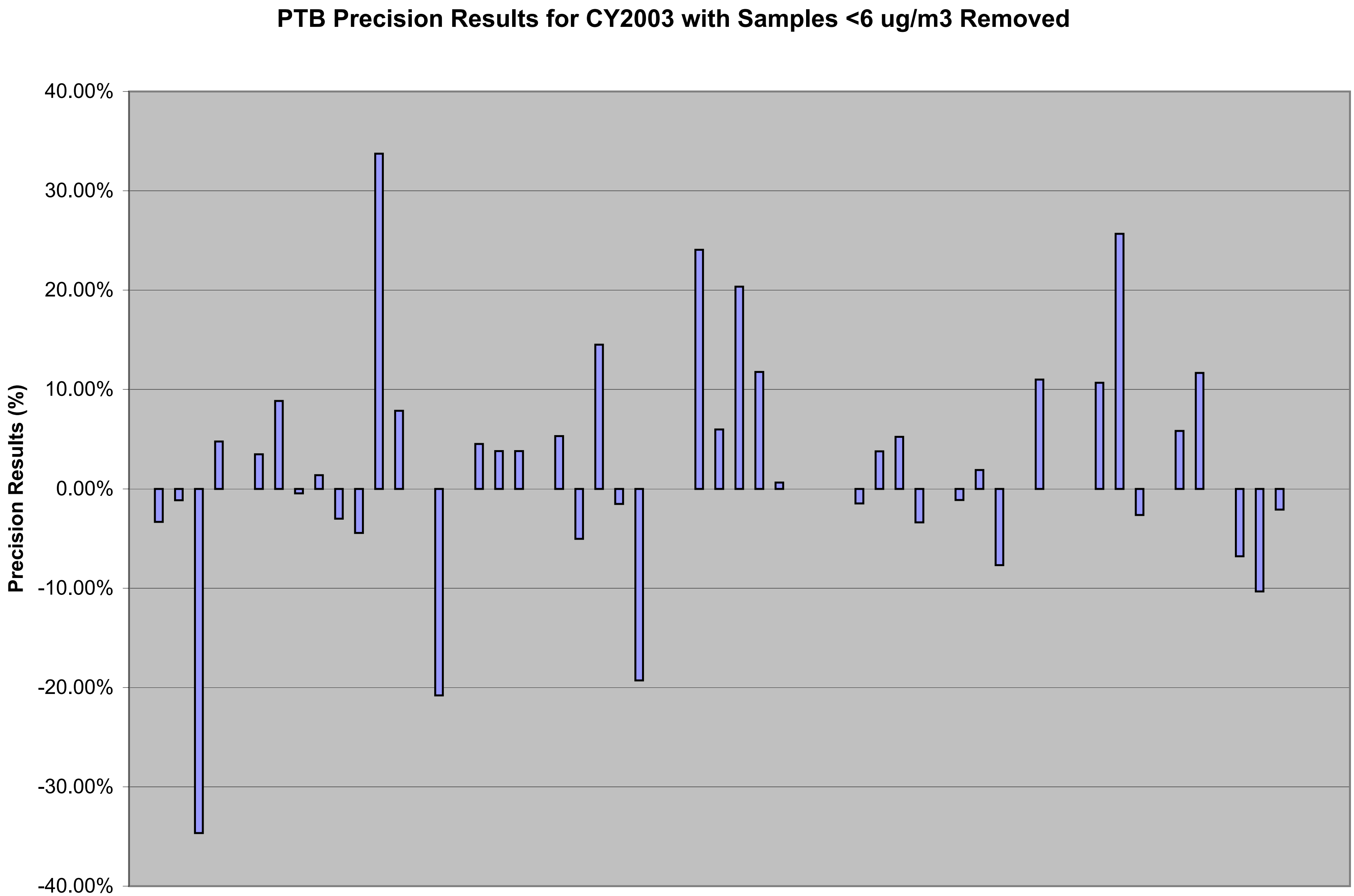
- Complete chemical analysis of archived filters; potentially use results to help explain observed sampler performance.
- Conduct detailed analysis of all data, including particle chemistry data and hourly performance of semi-continuous methods.
- Possibly perform laboratory tests with samplers to better understand aerosol fractionation and/or particle loss issues.
- Use study results as guidance during regulatory development of PMc testing requirements and acceptance criteria.

Acknowledgements

- Indiana Department of Environmental Management (Gary site)
- Maricopa County Environmental Services Department (Phoenix site)
- University of California-Riverside, Agricultural Operations (Riverside site)

Disclaimer

- This work has been funded in part by the United States Environmental Protection Agency under Contract 68-D-00-206 to ManTech Environmental Technology, Inc. It has been subjected to Agency review and approved for publication.
- Mention of trade names or commercial products does not constitute endorsement or recommendation for use.


PTB Precision Results for CY2003 with Samples <6 ug/m3 Removed
Precision Results (%)
40.00%
30.00%
20.00%
10.00%
0.00%
-10.00%
-20.00%
-30.00%
-40.00%

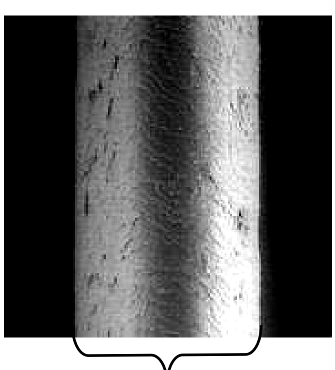
PM $_{10-2.5}$ (Coarse) Data Quality Objectives

EPA 23rd Annual National Conference on Managing Environmental Quality Systems

April 16, 2004

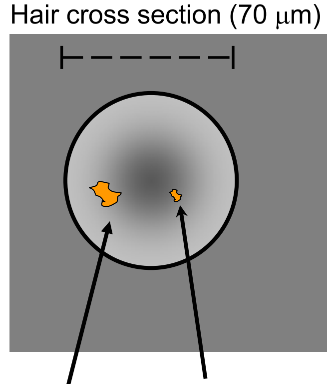
Particulate Matter: What is It?

A complex mixture of extremely small particles and liquid droplets



Human Hair (70 µm diameter)

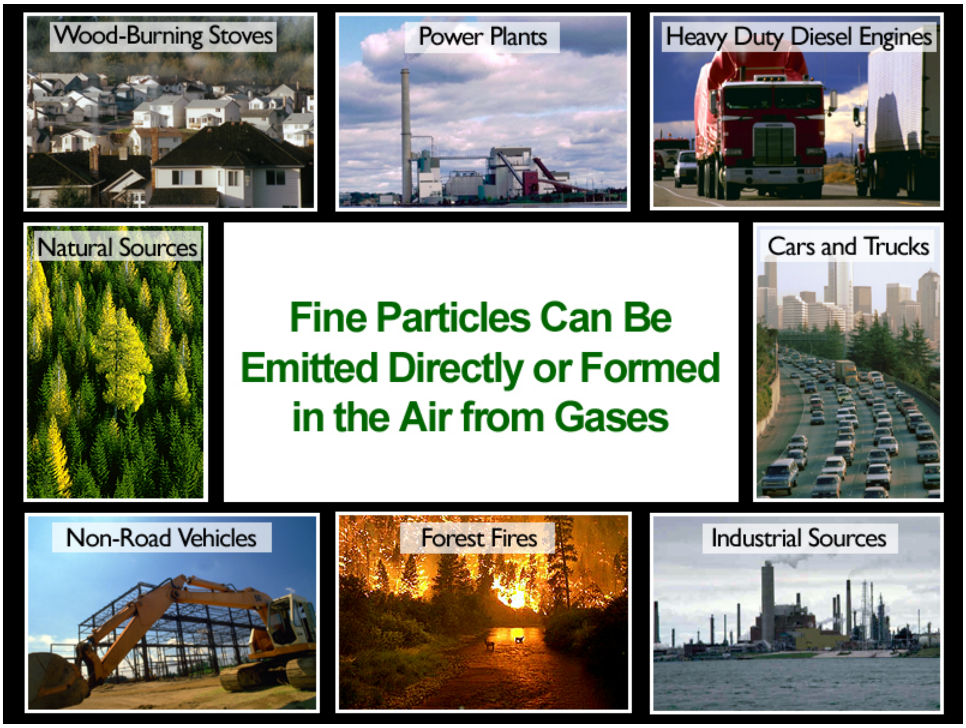
Hair cross section (70 µm)



PM_{10} (10µm)

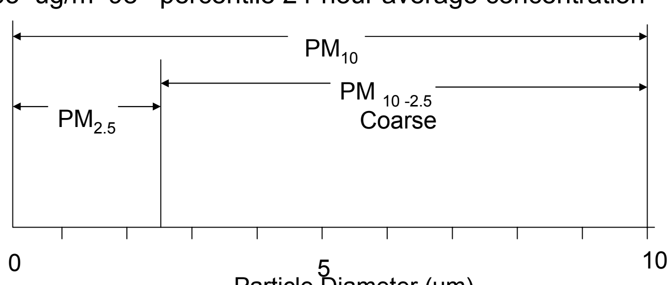
$PM_{2.5}$ (2.5 µm)

M. Lipsett, California Office of Environmental Health Hazard Assessment



National Ambient Air Quality Standards (NAAQS)

- **PM_{10}** –
  - 50 ug/m^3 annual arithmetic mean
  - 150 ug/m^3 24-hour average concentration
- **$PM_{2.5}$**
  - 15 ug/m^3 annual arithmetic mean
  - 65 ug/m^3 98th percentile 24-hour average concentration



What we'll be discussing

- Approach used to develop data quality objectives for $PM_{10-2.5}$
- Understanding the impacts of various sources of uncertainty
  - What sources of uncertainty have we studied?
  - Are certain sources more important than others?
- What are results specifically for levels of the standards proposed in the latest Staff Paper?
- What are implications of these results?
  - Identify appropriate gray zones for $PM_{10-2.5}$ leading to:
    - appropriate measurement quality objectives for $PM_{10-2.5}$
    - Appropriate methods for the program

Data Quality Objective Process

- Process to ensure that the data collected meet decision-maker needs.
  - The hardest part - finding out decision maker needs.
  - Once the needs are specified (and quantified) statistical models (simulation models in this case) can be used to quantify data quality or demonstrate how various data quality issues affect the quality of the end product.

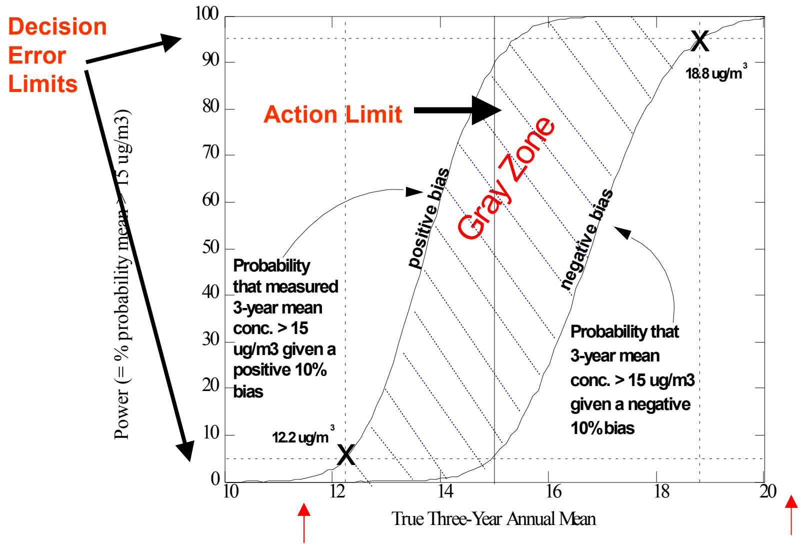
DQO Uses

- Initial planning- can be used to help understand data quality impacts on:
  - The level of the standard
  - The form of the standard
  - Decision errors
  - Network designs (number of sites, sampling frequency etc.)
- Ongoing monitoring implementation
  - Provides an excellent assessment tool for achievement of data quality
  - Provides a way to focus on quality system improvements as well as site specific improvements

DQOs

- Are not used to invalidate data
- Identifies the probability of decision errors not the fact that these errors have occurred.

Example Performance Curve



What variables can we vary to understand impact to performance curves?

- **Related to Standard**
  - Level of standard
  - Percentile for daily standard
  - Type I and II decision error rates
- **Related to Population**
  - Seasonality ratio
  - Population variability
  - Autocorrelation
- **Related to Measurement System**
  - Sampling frequency
  - Completeness
  - Measurement bias
  - Measurement precision

Issue- Two types of Methods Contemplated Continuous vs. Gravimeteric (Manual) Methods

- Continuous
  - Bias may be an issue in certain locations
  - May need to make them "FRM-like", will require 1 year of "calibration"
  - Improved sampling frequency and better completeness
- Gravimetric
  - Labor intensive (field/Lab)
  - More potential for measurement error and incomplete data
  - May be considered "truth"

Preliminary Precision Estimates From NERL Study

site/monitor	PMc n (runs used in precision calculation)	PMc number of complete runs	PMc precision
Gary_FRM	30	29	**5.4%**
Gary_R_P_Dichot	30	12	**4.2%**
Gary_Tisch	30	30	**10.3%**
Gary_TEOM	30	30	**2.9%**
Gary_APS	30	30	**24.5%**
Phoenix_FRM	29	18	**2.7%**
Phoenix_R_P_Dichot	20	5	**5.2%**
Phoenix_Tisch	30	30	**9.7%**
Phoenix_TEOM	30	30	**7.3%**
Phoenix_APS	20	20	**16.2%**
Riverside_FRM	30	29	**2.5%**
Riverside_R_P_Dichot	30	30	**1.8%**
Riverside_Tisch	30	30	**4.2%**
Riverside_TEOM	27	24	**4.5%**
Riverside_APS	20	20	**3.3%**
all_FRM	89	76	**4.1%**
all_R_P_Dichot	80	47	**3.8%**
all_Tisch	90	90	**10.1%**
all_TEOM	87	84	**6.2%**
all_APS	70	70	**19.5%**

Preliminary Bias Estimates using NERL Study Data

Bias Based on FRM 1 as Truth

PMc	All sites mean bias	Phoenix mean bias	Gary mean bias	Riverside mean bias
APS	**-45.0%**	-45.4%	-45.8%	-43.6%
APS_1	**-42.4%**	-44.1%	-37.9%	-46.3%
APS_2	**-47.8%**	-47.2%	-53.7%	-40.9%
Dichot	**-10.4%**	-18.8%	-9.4%	-6.5%
Dichot_1	**-12.1%**	-18.7%	-11.1%	-6.2%
Dichot_2	**-8.2%**	-15.1%	-9.3%	-6.2%
Dichot_3	**-10.8%**	-19.8%	-8.8%	-7.2%
FRM	**-1.0%**	1.2%	0.2%	-4.1%
FRM_2	**-4.7%**	-2.5%	-4.1%	-6.5%
FRM_3	**2.1%**	3.7%	4.4%	-1.7%
TEOM	**-16.9%**	7.0%	-31.0%	-26.2%
TEOM_1	**-17.8%**	10.9%	-32.8%	-30.0%
TEOM_2	**-20.8%**	0.0%	-32.7%	-30.6%
TEOM_3	**-12.1%**	10.1%	-27.7%	-18.0%
Tisch	**0.4%**	5.6%	-8.9%	5.1%
Tisch_1	**-1.8%**	3.1%	-15.6%	7.8%
Tisch_2	**2.3%**	14.1%	-5.9%	-0.7%
Tisch_3	**0.8%**	-0.3%	-5.3%	8.3%

Bias Based on Dichot 1 as Truth

PMC	All sites mean bias	Phoenix mean bias	Gary mean bias	Riverside mean bias
APS	**-33.1%**	-32.6%	-19.2%	-40.2%
APS_1	**-30.3%**	-31.2%	-2.0%	-43.0%
APS_2	**-36.2%**	-34.7%	-36.4%	-37.3%
Dichot	**0.9%**	1.2%	4.0%	-0.5%
Dichot_2	**2.3%**	9.3%	5.0%	0.1%
Dichot_3	**-0.2%**	-1.0%	2.9%	-1.0%
FRM	**13.4%**	24.4%	14.0%	3.8%
FRM_1	**14.7%**	23.6%	13.1%	6.8%
FRM_2	**7.9%**	20.2%	9.2%	-0.3%
FRM_3	**16.7%**	28.0%	19.4%	5.0%
TEOM	**1.0%**	31.4%	-22.9%	-21.1%
TEOM_1	**-0.2%**	35.8%	-24.4%	-25.2%
TEOM_2	**-3.9%**	23.0%	-24.7%	-26.0%
TEOM_3	**6.8%**	35.5%	-19.6%	-12.2%
Tisch	**18.0%**	30.2%	4.0%	11.9%
Tisch_1	**17.4%**	27.2%	0.8%	14.7%
Tisch_2	**19.4%**	40.4%	2.9%	5.7%
Tisch_3	**17.2%**	23.0%	8.3%	15.2%

Development of the $PM_{10\text{-}2.5}$ DQO

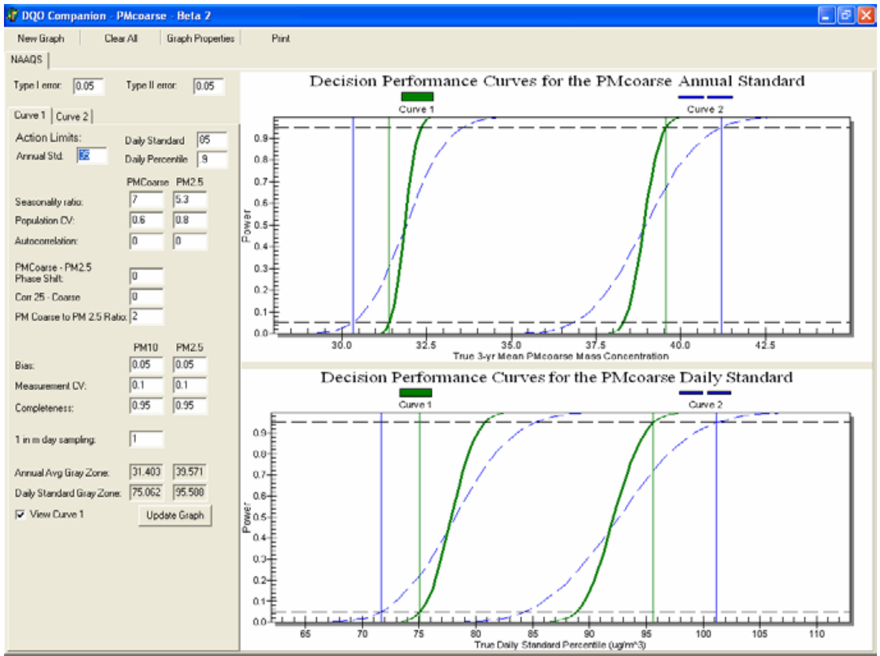
1. Aggregated all possible $PM_{10\text{-}2.5}$ data from information in AQS.
2. Reviewed data to identify the appropriate population distribution and population input parameters.
3. Reviewed measurement data quality indicators (precision, bias, completeness) and used default parameters from $PM_{2.5}$ DQO.
4. Calculate various decision performance curves depending on changes to input parameters

Approach

- Develop software that simulates data that are consistent with what is observed
- Simulation allows us to play "what if" games
- Simulations summarized as performance curves
  - For a range of possible true design values, curves tell how likely we are to observe a design value that is greater than the standard.
  - Performance curves developed for daily and annual standards

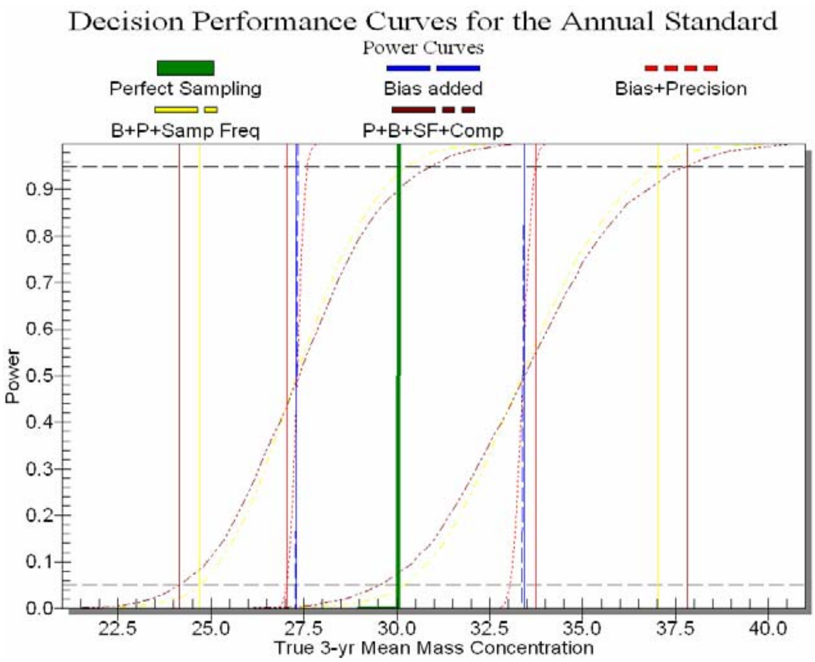
DQO Software Tools (Two Simulation Programs Available)

- Direct Measurements Tool - can be used to represent instruments that directly measure PM fraction of interest
  - Gravimetric methods (single filter) or continuous methods
- Difference-Method Measurements Tool - can be used to represent instruments that measure PM fraction of interest by using the difference between collocated instruments



Example Performance Curve

- Some Examples to Better Understand What Performance Curves Are and What They Show.
  - Perfect Sampling
  - Add Some Bias
  - Add Some Measurement Error
  - Sample Infrequently



Performance Curve Results for PMCoarse Levels of Standards Under Consideration
(Sampling frequency every day unless noted as 1/6)

- Annual Std 30
  - Direct: [26.7, 34.1]
  - Difference: [24.4, 38.3]
  - Difference:[22.9 , 41.6] 1/6
- Daily Std 75
  - Direct: [62.6, 88.5]
  - Difference: [59.7, 95.7]
  - Difference: [43.1, 102.9] 1/6
- Annual Std 13
  - Direct: [11.6, 14.8]
  - Difference: [10.6, 16.6]
  - Difference: [9.9, 18.0] 1/6
- Daily Std 30
  - Direct: [25.3, 35.7]
  - Difference: [24.2, 38.7]
  - Difference: [17.6 , 41.9] 1/6

(Based on daily sampling, 98th percentile, coarse seasonality of 7, 2.5 seasonality of 5.3, coarse population cv of 0.6, 2.5 popn cv of 0.8, coarse to 2.5 ratio of 2.25, autocorrelation of 0, bias=10%, meas cv=10%, 75% completeness.)

Which variables have larger impact on performance curves?

- **Related to Standard**
  - Level of standard
  - Percentile for daily standard
  - Type I and II decision error rates
- **Related to Population**
  - Seasonality ratio
  - Population variability (depends on sampling frequency)
  - Autocorrelation
- **Related to Measurement System**
  - Sampling frequency
  - Completeness
  - Measurement bias
  - Measurement precision (for daily standards)

Preliminary Assessment

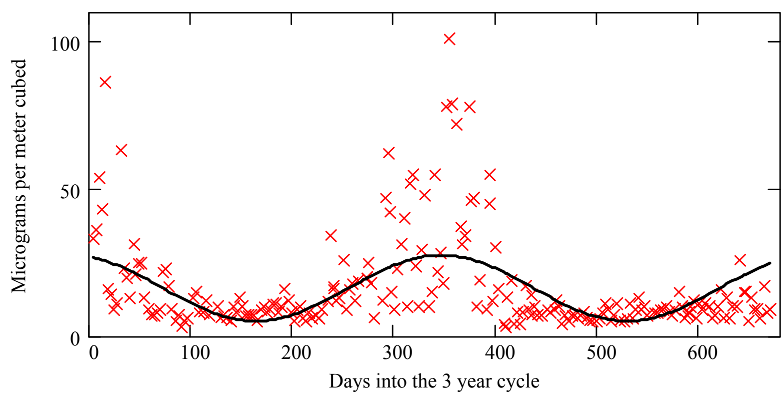
- Every day sampling with continuous monitoring is the best way of tightening the gray zone without changing measurement uncertainty input parameters.
  - However, what's considered truth; bias has to be understood and controlled

DQO Issues

- Do we shift gray zones to protect health?
- Do we keep decision errors on both sides of the action limit symmetric?
- How tight do we need the gray zones?
- Where do we get our bias estimate?

Back-up Slides

Example of One Simulated Time Series of PM2.5



Thousands of such simulations are generated and summarized in performance curves.

Impacts

- Example Impact of Sampling Frequency
  - 1 in 1: [13.3, 17.2] & [53.6, 77.9]
  - 1 in 3: [12.6, 18.2] & [43.6, 83.9]
  - 1 in 6: [12.1, 19.0] & [33.6, 84.8]
- Example Impact of Data Completeness
  - 75%: [12.1, 19.0] & [33.6, 84.8]
  - 85%: [12.2, 18.8] & [50.9, 103.5] Huh???

Impacts (continued)

- Example Impact of Bias
  - 5%: [12.7, 18.0] & [34.8, 81.3]
  - 10%: [12.1, 19.0] & [33.6, 84.8]
  - 20%: [11.1, 21.3] & [30.8, 96.0]
- Example Impact of Measurement Precision
  - 10% CV: [12.1, 19.0] & [33.6, 84.8]
  - 20% CV: [12.0, 19.2] & [32.4, 84.5]
- Example Impact of Population Variability
  - 80% CV: [12.1, 19.0] & [33.6, 84.8]
  - 50% CV: [12.6, 18.1] & [40.6, 81.5]

Issues- Continuous vs. Gravimetric (manual) Methods

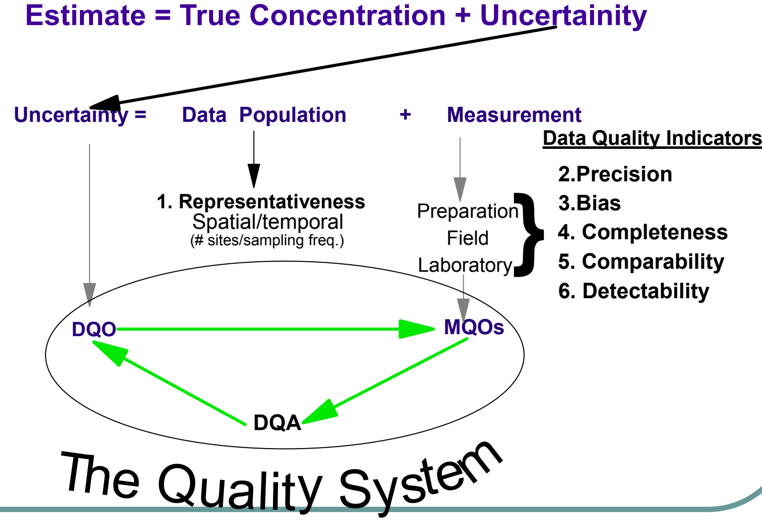
- Continuous
  - Bias may be an issue in certain locations
  - May need to make them "FRM-like", will require 1 year of "calibration"
  - Improved sampling frequency and better completeness
- Gravimetric
  - Labor intensive (field/Lab)
  - More potential for measurement error and incomplete data
  - May be considered "truth"

Performance Curve Results for PM2.5 Levels of Standards Under Consideration

- Annual Std 15
  - 1 in 1: [13.2, 17.2]
  - 1 in 6: [12.2, 18.8]
- Daily Std 50
  - 1 in 1: [41.1, 59.5]
  - 1 in 6: [25.4, 65.5]
- Annual Std 12
  - 1 in 1: [10.6, 13.8]
  - 1 in 6: [9.6, 15.2]
- Daily Std 30
  - 1 in 1: [24.7, 36.1]
  - 1 in 6: [15.2, 39.6]

(Based on 98th percentile, 2.5 seasonality of 5.3, 2.5 popn cv of 0.8, autocorrelation of 0, bias=10%, meas cv=10%, 75% completeness.)

Estimate = True Concentration + Uncertainity



New Quality Indicator Statistics for the Gaseous Criteria Pollutants

Presented at:

EPA 23rd Annual Conference on Managing Environmental Quality Systems

April 13-16, 2004

Tampa, Florida

AMBIENT AIR SESSION III

Thursday, April 15, 2004 - 1:30 pm

Presented by: Basil Coutant
Battelle, 505 King Avenue, Columbus, OH 43201
coutantb@battelle.org

Acknowledgements

- Shelly Eberly, EPA/ORD
- Mike Papp, EPA/OAQPS
- Chris Holloman, Battelle
- Kristen Swinton, Battelle

Work supported by U.S. EPA/OAQPS

EPA Contract No. 68-D-02-061

Overview

- Goals
- Review current data and statistics
- Present the new statistics
- Discussion
- Conclusions

Purpose / Goal

- The goal for this project was to develop a set of statistics that correspond to the DQO statements for the gaseous pollutants.
- The statistics need to be based on the quality indicator data already collected.
- If possible, it was desired to use statistics that would be consistent with the PM2.5 program.

The Data and Current Statistics

Pollutant	Data	Current Summary Statistic(s)
NO_2, SO_2, CO, and O_3	Biweekly Precision Checks Annual Accuracy Audits	Probability Interval Probability Interval None
Lead	Flow rate audits Lead strip audits Co-located measurements	Probability Interval Probability Interval Probability Interval
PM_{10}	Flow rate audits Co-located measurements	Probability Interval Probability Interval
$PM_{2.5}$	Flow rate audits Co-located measurements PEP measurements	Mean percent deviation from target flow. CV estimated using the Root-Mean Square of percent differences. Bias estimated by mean of percent differences.

The Probability Interval

- For the gaseous pollutants, the main quality assurance tools are the “biweekly” precision checks. The precision checks are made by sampling from air with a known concentration of a given pollutant.
- A probability interval based on the relative percent error of these checks is created. This probability interval is the main method for summarizing the relative percent errors and serves well as a summary tool.
- It does not control precision and bias separately, and these two do not contribute equally to decision errors.

"Biweekly Precision" Data

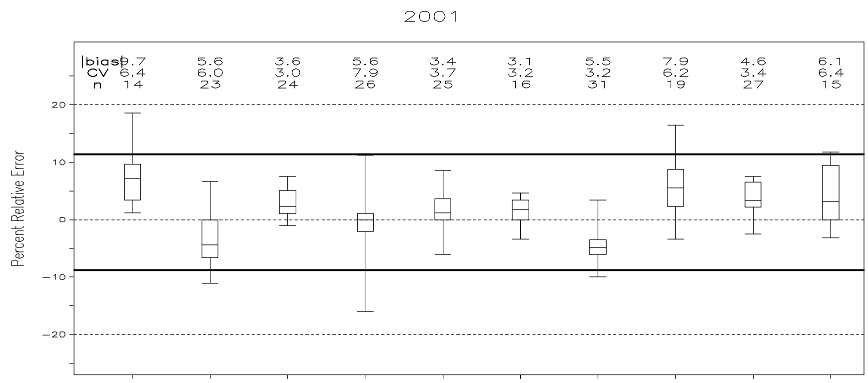
- Repeated measurements against the same "truth."
- These can be used to measure both precision and bias.
- Moreover, for automated methods they are frequently not biweekly.
- I suggest the name "Single-Point Checks".

	CO	NO_2	O_3	SO_2
Single-point check range (PPM)	**8-10**	**0.08-0.10**	**0.08-0.10**	**0.08-0.10**

Do we want to estimate?

- No! We want to control the bias and precision.
  - We do not want the best possible statistical estimates.
  - Instead we want summary statistics that let us know whether or not bias and precision are being controlled at the site level, even though the statistics are often at the reporting agency level or higher.
  - We also want to capture any season variations without allowing cancellation.

Example (Ozone 2001 Data)



The Statistics

For each single-point check, calculate the relative percent error, *d*,

$$d = \frac{ind - act}{act} \cdot 100$$

where *ind* is the concentration indicated by the agency’s measurement and *act* is the actual concentration being measured.

Bias

The bias statistic is an upper bound on the mean absolute values of the relative errors.

$$|bias| = AB + t_{0.95,n-1} \cdot \frac{AS}{\sqrt{n}}$$

where:

- n is the number of single-point checks being aggregated;
- $t_{0.95,n-1}$ is the 95th quantile of a t-distribution with n-1 degrees of freedom;
- AB is the mean of the absolute values of the d's; and
- AS is the standard deviation of the absolute values.

Bias (cont.)

In particular, AB and AS are:

$$AB = \frac{1}{n} \cdot \sum_{i=1}^{n} |d_i|$$

$$AS = \sqrt{\frac{n \cdot \sum_{i=1}^{n} |d_i|^2 - \left(\sum_{i=1}^{n} |d_i| \right)^2}{n(n-1)}}$$

Precision

The precision statistic is an upper bound on the standard deviation of the relative errors.

$$CV = \sqrt{\frac{n \cdot \sum_{i=1}^{n} d_i^{\,2} - \left(\sum_{i=1}^{n} d_i\right)^2}{n(n-1)}} \cdot \sqrt{\frac{n-1}{c_{0.05,n-1}}}$$

where $c_{0.05,n-1}$ is the 5th percentile of a chi-squared distribution with n-1 degrees of freedom.

Verifying Assumptions

- The accuracy audits are annual NIST traceable audits over a range of concentrations.
- These accuracy audits can be used to verify the results obtained from the single-point checks and to validate those results across a range of concentration levels.
- Annual and three-year agency-level probability limits calculated from all the *single-point checks* should capture approximately 95 percent of the relative percent differences from the accuracy audits (for all levels).

Probability Interval

The current probability limit statistics should be kept for the single-point checks, but compared to the accuracy audits.

$$Upper\ probability\ Limit = m + 1.96 \cdot S$$

$$Lower\ probability\ Limit = m - 1.96 \cdot S$$

$$m = \frac{1}{k}\sum_{i=1}^{k} d_i \qquad S = \sqrt{\frac{k \cdot \sum_{i=1}^{k} d_i^{\,2} - \left(\sum_{i=1}^{k} d_i\right)^2}{k(k-1)}}$$

Discussion - Bias

- Since the bias is the more influential of the two types of error on decision quality, the bias is the more strongly controlled under the scheme.
- The bias statistic has two conservative components:
  - The absolute values were chosen to detect or control for cases where the bias is positive part of the time and negative part of the time.
  - The use of a confidence limit upper bound adds an additional protection, in this case, against random errors in the estimate of the mean of the absolute relative errors.
- Neither of the above is consistent with the PM2.5 program, but both are being considered for the PM2.5 program.

Discussion - Precision

- The confidence limit upper bound protects against random errors in the estimate of the standard deviation.
- The DQO quantity of interest is the CV of the measurement error, so it would not be appropriate to use the standard deviation of the absolute values as in the bias statistic.
- The statistic is less conservative than the root-mean-square statistic currently used for PM2.5, because it includes a mean correction (the second term under the first square root). This was felt to be appropriate for the gaseous pollutants.
- Moreover, the precision statistic is being considered as a replacement for the current statistic used for precision in the PM2.5 program.

Discussion - Accuracy Audits

- The accuracy audits tie everything together.
- There are not enough data from these to get summary information from them alone.
- Instead, they are consistency and assumption checks under the proposed scheme.

Conclusion

- The statistics presented make better use of the QA data currently collected by the State and Local agencies monitoring the gaseous criteria pollutants.
- They separately control the precision and bias as required by the DQO statements.
- They are not estimates of precision and bias, but rather upper bounds to control the bias and precision.
- They incorporate both the single-point check data and the accuracy audit information.

EPA Infrastructure for Ambient Air Bias Traceability to NIST

Changes and Status

Mark Shanis, OAQPS

Tampa, 2004

GOAL:STRONGER REGIONAL SUPPORT

FOCUS ON 3 PROGRAMS

EPA NPAP: MAILED (?PSD) + MOBILE TTP

SRP: 2 UPGRADES, BASE IN LV

PROTOCOL GASES VERIFICATION: 3RD PARTY

NPAP(M+TTP) + PEP = NPEP

- 2003 TRANSITION: Mailed Only, Back of the Analyzer (BOA) to Mailed + Mobile Through-the-Probe (TTP)/Station Inlet
- 2003 Audits:3 Regions did TTP audits,1/3 did PEP+TTP; Mailed to R1,2,3,8,9,and 10
- May 2004:1st group training and certification, Like PEP (Written+ Hands-on): R2,4-7,9
- SOP (Adding to Draft as Use) and Implementation Plan (still in Prep).
- New:Reg2 Hi Flow Rate Subsystem (May-June)

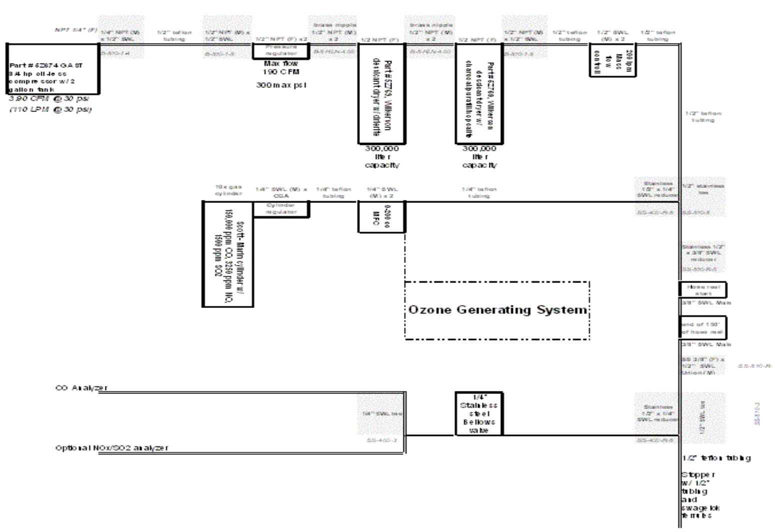
April 2004 NPAP Status

- Mobile TTP Labs in 6 Regions
- Del'd. to 2&9 on 3/11&12; 4 in Mid-Jan
- 2/6 Will Share Lab in 3 Neighboring Regions (9+10;2+1&3);R8?- now asking
- All 6 Plan Audits in 2004;So far-
  - R6= 17 sites, 26 analyzers, from 8/03- 2/04
  - R7= 16 sites,19 analyzers, from 7/03-3/04
  - R4&9; R5= Field Audits start in April; May

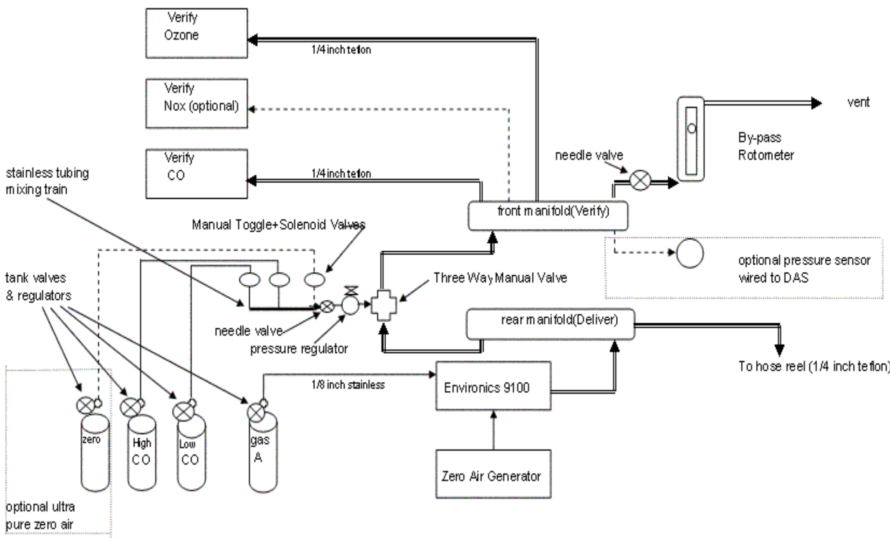
Combination Benefits/TradeOffs

- If EPA Staff Do Audits: No ESAT Labor Costs; In Practice So Far: R7-Can Focus on TTP
- If ESAT Staff Do Audits: Only EPA trainng, Oversight Time Needed; R6 1st, soon 4,9
- Common Benefit: Expanded Resources, Can Do Both TSAs and PEs more easily
- 04:EPA 2,5,7,TTPonly;ESATR4,6,9,NPEP

R2 High Flow Generation System



CAN 1 PERSON DO IT?



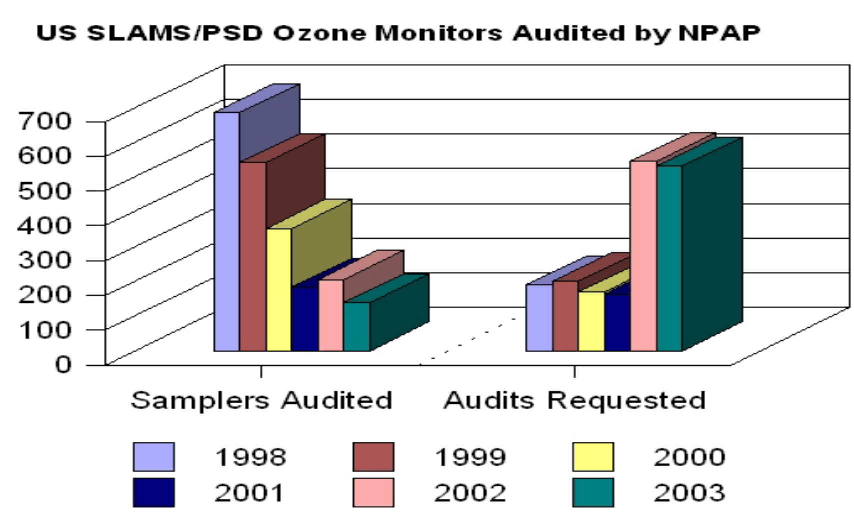
Cost/Benefit Tradeoffs

- Reg 6 NPEP: 17 sites, 26 monitors-$29K
- Amounts for ESAT in 04
  - R2: $10K, mostly for trning, support EPA TTP
  - R4:$35K, for trning, PEP+TTP (NPEP)
  - R5, $35K, trng, support EPA TTP, some TTP
  - R6: $35K, for NPEP (PEP+TTP)
  - R7: $25K,Trng, Support EPA TTP
  - R9: $30K,Trng, PEP+TTP

Mailed Program, Reductions

- Mailed Contractor Funded in 04' for Regions 1,2,3,8,9, and 10
- Provide in 04': Ozone, CO, SO2, NO/NO2, PM10, Lead
- Probably not:PAMS-VOC, Carbonyls; PSD
- Max# Audit Mailed Devices Very Limited: 15-20 O3, 5-6 CO/SO2/NO/NO2;10 PM10
- Next Yr Funding? IMPT: S&L OK 103 use

Example: Ozone Monitors, NPAP Audits, 1998-2003



NPEP Summary

- Special Advantages:
  - Multi-functionality: Audits, Sampling Priorities
  - Enhanced Equipment and labor, Regionally-Based, making it easier to do PEs and TSAs, and High Priority Sampling, Training, and Support for New Methods for S&Ls in Region
- Issues: Funding Future?
  - Cost/Audit Result Data for S&L 103 OK in'05
  - Hi Flow Stations, Portability; PSD, PAMS; Data Base System for Past NPAP and Future

Region 7 Mobile TTP Audit Lab

Roof Platform+Sampling Mods.

13

TEOM Mod +PC,etc.

14

Interior Front, Sampling Mod.

15

EPA/NIST SRP Network

- STATUS
  - In Regions 1, 2, 4, 5, 6, 7, 8 and 9; 2 does 3, 9 does 10; a 2nd in 9 approx. this July
  - 2 originally set up for comparing the 8 Regional SRPS to NIST,1 traveling, 1stationary;1st based in RTP; now in LV
  - Range of Ages:1st RTP, Done 2/83, last in KC, KS(R7)1/89
  - NIST has 12 Worldwide, latest made this year

16

EPA SRP Network Changes

- Upgrades Needed: 2 Funded, 2nd still in process; Hardware and software
- Feature Improvements: Change from all Manual operation to ability to automatically perform and record required documented procedure
- Benefits- Easier to certify multiple primary or transfer standards, more consistently, and with lower zero signal
- ORIA-LV Plans: Improved Trouble-shooting; Grp OK to Std. Cert. Forms; Summary Reports

NIST SRP Network Changes

- NIST Talk at June AWMA mtg-will provide first documentation of international comparisons, including EPA network
- Plans in progress to have BIPM(France) Lead as European Center and for non-USA SRP support
- Cost of new SRP Rising (Approx. $65K now); Revised Manual in Progress

EPA Traceability Protocol

- On EPA TTN/EMC; for source and ambient levels, as of 98
- Presentation at this meeting
- ORD Verification Program stopped mid90s
- Users reporting problems; EPRI and EPA have done studies recently to assess problems
- Some Vendors have requested restart of Verification

3rd Party Verification

- Critical Features of ORD Audit Program
  – Low Cost
  – Very Low number of samples
  – Audit Sample Buying unknown to Vendor
  – Experienced Lab analysis; Vendors Coded
  – Process independent of vendors
  – RESULTS REPORTED TO PUBLIC
  – Documented Improvement in Tag Accuracy

Bias Traceability Summary

- Programs are still active, changes occurring; Cited in Proposed 40 CFR Part 58 Revision
- Quality Data Requires both Continuation of Support and Change to keep up with Method and Data Priority changes
- Protocol Verification Success Indicates High sample numbers are not the only determinant of Effect:
  - Users Respond to the Attention Brought by Open Bias Assessment

2003 Blind Audit of EPA Protocol Gases

John Schakenbach, U.S. EPA, CAMD

Bob Wright, U.S. EPA, ORD

Joe Elkins, U.S. EPA, OAQPS

Scott Shanklin, Cadmus Group

April 15, 2004

Why are we giving this talk to a QA Audience ?

- EPA Protocol Gases are widely used gaseous reference standards
- QA professionals need to understand the uncertainty of these reference standards
- This program is a example of how EPA can assess a commercial product with minimal interference and reasonable cost
- Useful lessons about organizing an audit program and about gas metrology

Characteristics of EPA Protocol Gases

- They must be traceable to NIST reference standards (Standard Reference Materials)
- Anyone may use the protocol to certify compressed gas mixtures(vendors, users, gov't)
- A general, flexible analytical procedure
- A specific statistical analysis procedure
- Specific documentation requirements
- EPA conducts audits to determine their accuracy

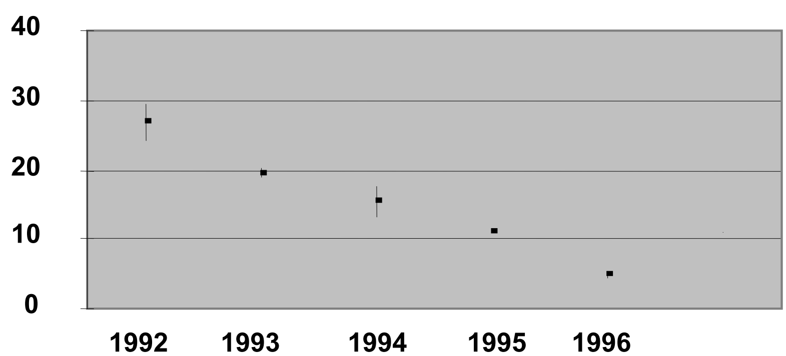
Why is there a need for the EPA Protocol Gas Audit Program ?

- EPA does not certify or permit specific organizations to produce these standards. Anyone can do so.
- EPA does not inspect or audit vendor facilities
- The protocol is a general analytical procedure. The analyst chooses specific procedures and then calculates the uncertainty of the measurements.
- The protocol does not have an acceptance criterion for the uncertainty of standards. The user specifies it. The Acid Rain Program specifies +/- 2% accuracy.
- The audits are the only tool available for EPA to obtain an independent assessment of the uncertainty.

History of EPA Audit Program

- From 1985 to 1997, there were 253 audits
- 78% of standards accurate to within +/- 2%
  95% of standards accurate to within +/- 5%
  99% of standards accurate to within +/- 10%
- In 1995, one cylinder biased by -16.3%
- Strong utility and vendor support for audits
- Audit Program ended in 1998

Audits are strongly correlated with improved quality



Percentage not meeting acceptance criterion

2003 Audit of EPA Protocol Gases

- First audit in 7 years
- Blind audit (vendors didn't know)
- 14 national specialty gas vendors
- 42 tri-blend cylinders (3 per vendor)
- Similar audit procedures as in past
- SRMs and NTRMs used as reference stds.
- Mactec (primary audit lab) and
  Spectral Insights (reference audit lab)

Tri-blend EPA Protocol Gases

	CO_2 (%)	NO (ppm)	SO_2 (ppm)
Low	5	50	50
Medium	12	400	500
High	18	900	1000

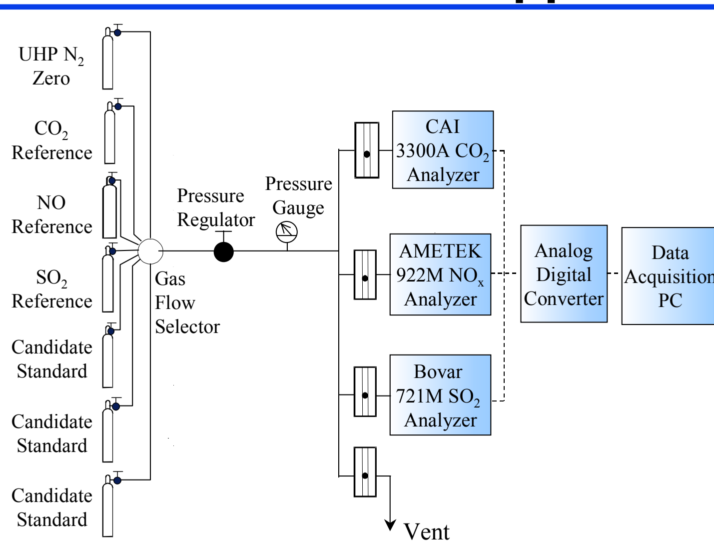
Analytical Instrumentation

- NO - API Model 200AH chemiluminescence
- NO - Ametek Model 922M UV absorption
- SO_2 - Bovar Model 721M UV absorption
- CO_2 - California Analytical Model 3300A NDIR
- NO, SO_2, and CO_2 - Nicolet Model 760 FTIR
- Environics Series 3740 gas dilution system

Mactec Lab

Instrumentation

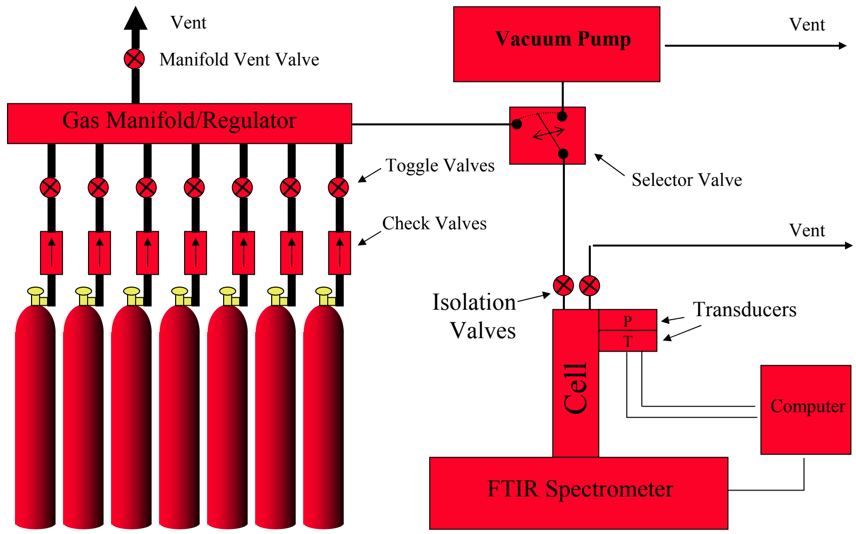
Schematic of Mactec Apparatus



Spectral Insights Mobile FTIR Lab

Nicolet Nexus Model 760 FTIR

Spectral Insights Apparatus



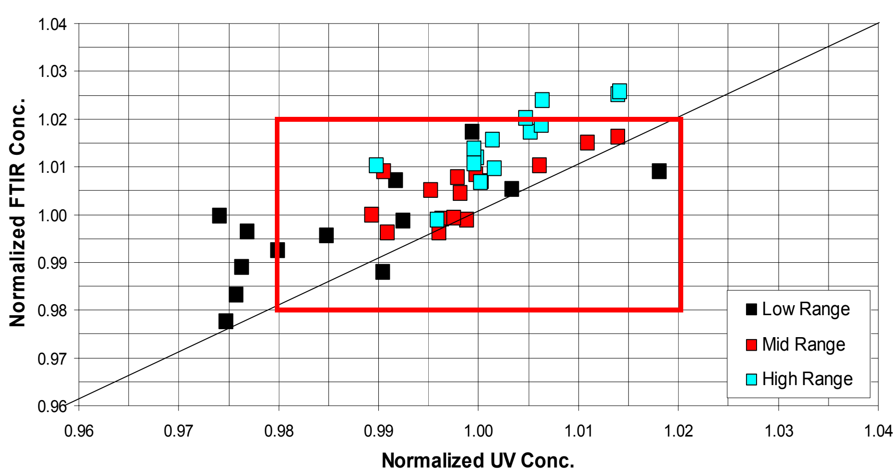
Instrumentation Problems

- High-level CO_2 SRM empty for FTIR analyses
  - FTIR lab prepared a high-level CO_2 primary ref. std.
  - EPA threw out the high-level CO_2 FTIR data
- NO data from chemiluminescent analyzer biased low due to CO_2 quenching
  - Chemiluminescent NO data thrown out
  - Measurements repeated with a NO UV analyzer

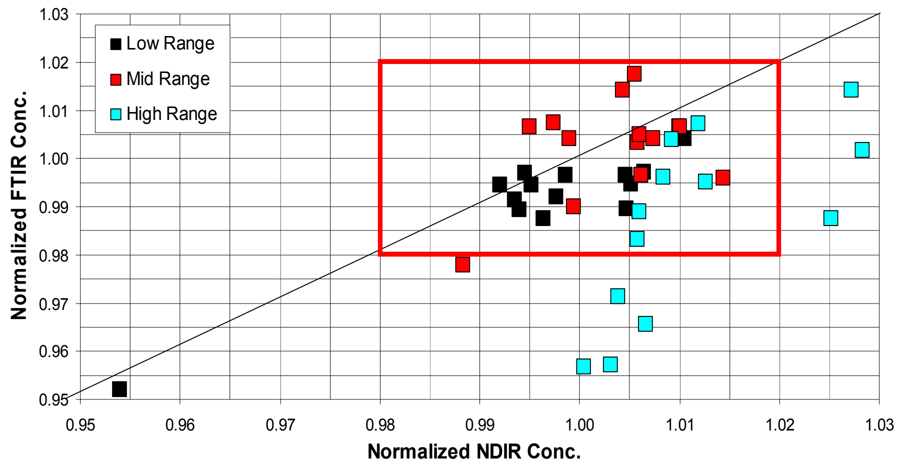
Instrumentation Problems

- NO UV analyzer set up for 0 - 500 ppm range, but should have been for 0-1000 ppm range
- SO_2 interfered with NO UV analyzer readings
  - Injected SO_2 in N_2 mixture to develop a interference correction equation for NO data, but curve for SO_2 and NO in N_2 mixture is very nonlinear at mid- and high-level concentrations
  - EPA threw out mid- and high-level NO UV data

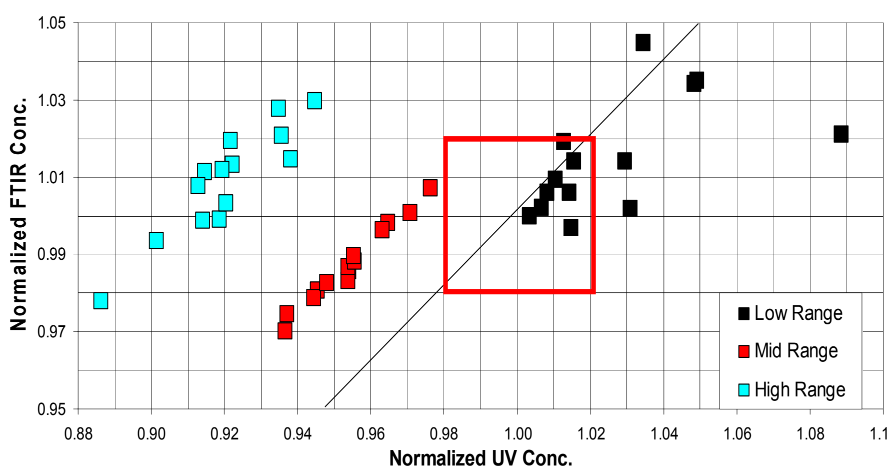
Comparison of FTIR Data with UV Data for SO_2 EPA Protocol Gases



Comparison of FTIR Data with NDIR Data for CO_2 EPA Protocol Gases



Comparison of FTIR Data with UV Data for NO EPA Protocol Gases



EPA Protocol Gases not meeting Acid Rain Program's Acceptance Criterion for One or Both Audit Analyses

	NO Analyses			SO_2 Analyses			CO_2 Analyses		
	UV	FTIR	Both	UV	FTIR	Both	NDIR	FTIR	Both
Low	6/14	4/14	4/14	6/14	1/14	1/14	1/14	1/14	1/14
Mid	---	3/14	---	0/14	0/14	0/14	0/14	1/4	0/14
High	---	3/14	---	0/14	3/14	0/14	3/14	- - -	- - -

Summary of Results

- Overall failure rate: 32 of 210 analyses (15%)
- SO_2 failure rate: 10 of 84 analyses (12%), worst bias 2.7%
- NO failure rate: 16 of 56 analyses (29%), worst bias –8.4%
- CO_2 failure rate: 6 of 70 analyses (9%), worst bias 5%
- All documentation requirements were met

Lessons Learned for Future Audits of EPA Protocol Gases

- Detailed audit SOPs are needed
- Audit labs need experience in gas metrology
- Instrumentation must be modified for gas metrology
- Traceability protocol needs to be modified for FTIR
- Gain experience with single component mixtures before moving to multicomponent mixtures
- Check multicomponent interference effects beforehand
- Intercompare audit labs before the audit starts
- Use an SRM or NTRM for FTIR measurements

Protocol Gas Audit Program Direction

- Scope
- Structure
- Funding
- Oversight
- Protocol Revision and Updates

Next Steps

- Detailed Outline
- Get feedback